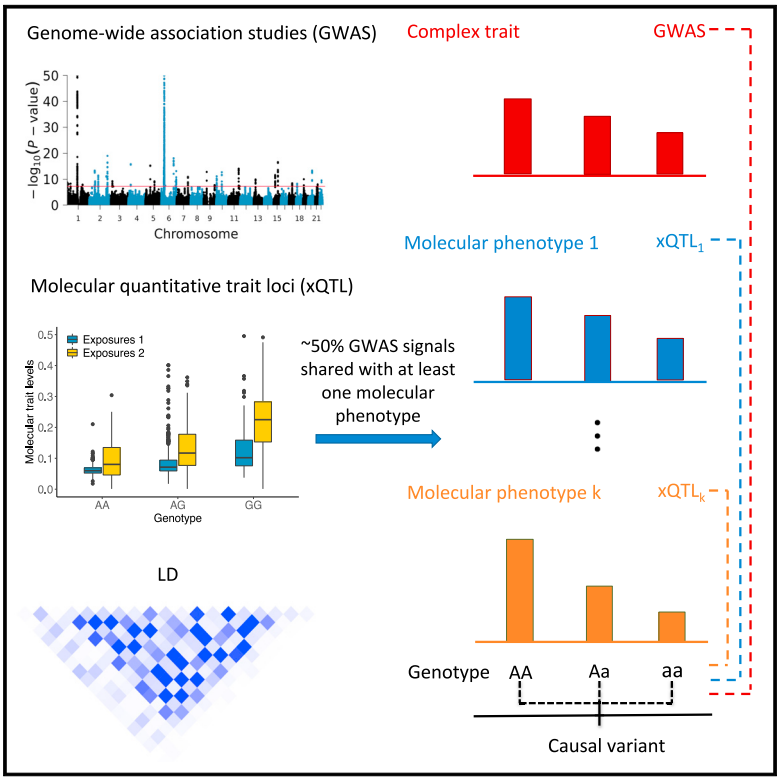
Cell Genomics

Article

Joint analysis of GWAS and multi-omics QTL summary statistics reveals a large fraction of GWAS signals shared with molecular phenotypes

Graphical abstract



Highlights

- Joint analysis of GWAS and multi-omics xQTL data enhances complex trait gene discovery
- Approximately 50% of GWAS signals are shared with at least one molecular phenotype
- New insights into the genetic regulatory mechanisms underlying complex traits

Authors

Yang Wu, Ting Qi, Naomi R. Wray, Peter M. Visscher, Jian Zeng, Jian Yang

Correspondence

j.zeng@uq.edu.au (J.Z.), jian.yang@westlake.edu.cn (J.Y.)

In brief

Wu and colleagues developed OPERA, a method that integrates genetic analysis of multi-omics data with GWASs to explore the biology of genetic associations. They found that 50% of GWAS signals are shared with at least one omic measurement. This approach may aid in comprehending the underlying mechanisms of complex traits.

Wu et al., 2023, Cell Genomics 3, 100344
August 9, 2023
https://doi.org/10.1016/j.xgen.2023.100344

CellPress

Cell Genomics

CellPress
OPEN ACCESS

Article

Joint analysis of GWAS and multi-omics QTL summary statistics reveals a large fraction of GWAS signals shared with molecular phenotypes

Yang Wu,[1] Ting Qi,[2,3] Naomi R. Wray,[1,4] Peter M. Visscher,[1] Jian Zeng,[1,5,*] and Jian Yang[2,3,5,6,*]
[1]Institute for Molecular Bioscience, The University of Queensland, Brisbane, QLD 4072, Australia
[2]School of Life Sciences, Westlake University, Hangzhou, Zhejiang 310024, China
[3]Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou, Zhejiang 310024, China
[4]Queensland Brain Institute, The University of Queensland, Brisbane, QLD 4072, Australia
[5]These authors contributed equally
[6]Lead contact
*Correspondence: j.zeng@uq.edu.au (J.Z.), jian.yang@westlake.edu.cn (J.Y.)
https://doi.org/10.1016/j.xgen.2023.100344

SUMMARY

Molecular quantitative trait loci (xQTLs) are often harnessed to prioritize genes or functional elements underpinning variant-trait associations identified from genome-wide association studies (GWASs). Here, we introduce OPERA, a method that jointly analyzes GWAS and multi-omics xQTL summary statistics to enhance the identification of molecular phenotypes associated with complex traits through shared causal variants. Applying OPERA to summary-level GWAS data for 50 complex traits (n = 20,833–766,345) and xQTL data from seven omics layers (n = 100–31,684) reveals that 50% of the GWAS signals are shared with at least one molecular phenotype. GWAS signals shared with multiple molecular phenotypes, such as those at the *MSMB* locus for prostate cancer, are particularly informative for understanding the genetic regulatory mechanisms underlying complex traits. Future studies with more molecular phenotypes, measured considering spatiotemporal effects in larger samples, are required to obtain a more saturated map linking molecular intermediates to GWAS signals.

INTRODUCTION

Over the past decade, genome-wide association studies (GWASs) have identified hundreds of thousands of genetic variants associated with a wide range of complex traits (including diseases) in humans.[1] However, the biological mechanisms underlying the variant-trait associations are mostly unknown. Given that the trait-associated variants often appear in genomic regions enriched with regulatory elements and molecular quantitative trait loci (xQTLs),[2–7] one approach to understand the mechanisms that derive the associations is to explore whether the molecular phenotypes, such as transcription or protein abundance of a gene, are associated with the complex traits due to pleiotropy.

Different methods have been proposed to detect associations between molecular phenotypes and complex traits by integrating GWAS and xQTL data.[6,8–14] One class of methods imputes unobserved measurements of molecular phenotypes for individuals in a GWAS cohort and tests for associations between the imputed measurements of molecular phenotypes and a complex trait of interest (e.g., PrediXcan and TWAS).[9,11] Another class of methods detects associations between molecular phenotype measurements and the trait by assessing the concordance of GWAS and xQTL signals (e.g., COLOC and RTC).[6,10,12–15] These methods typically involve a metric to assess whether the overlap of GWAS and xQTL signals at a genomic locus is due to shared causal variant(s). Summary-data-based Mendelian randomization (SMR)[8,16] is a method that belongs to the second class. It uses the Mendelian randomization (MR) method framework to estimate the effect of a molecular phenotype (x) on the trait (y), followed by the heterogeneity in dependent instruments (HEIDI) test that utilizes linkage disequilibrium (LD) pattern to distinguish pleiotropy (x and y share the same set of causal variants) from linkage (causal variants for x are distinct from, but in LD with, those for y). By applying these methods to GWASs with gene expression QTL (eQTL) or protein QTL (pQTL) data, many genes have been prioritized as responsible for GWAS signals.[8–11,16,17] Integrating other types of xQTL data, such as DNA methylation (DNAm) QTL (mQTL), histone modification QTL (hQTL), and splicing QTL (sQTL), has provided mechanistic insights at some of the GWAS loci.[4,16,18,19]

The aforementioned methods only allow integration of GWAS data with xQTL data for one type of molecular phenotype at a time. When multiple types of xQTL data are available, analyzing each of them separately with GWAS data and testing all pairwise associations between the xQTL datasets to understand connections among the molecular phenotypes[16] is cumbersome and increases the multiple testing burden when controlling for

genome-wide type I error rate. Multiple trait colocalization (MOLOC) is an approach that allows colocalization analysis of xQTL summary data from multiple molecular phenotypes with GWAS data.[20] However, such analysis becomes computationally challenging when involving over three molecular phenotypes because the number of hypotheses to be tested increases exponentially with the number of molecular phenotypes involved.[21] To circumvent the computational hurdle, an efficient deterministic Bayesian approach (i.e., HyPrColoc) was developed to analyze a large number of traits using an approximate posterior probability of full colocalization.[21] However, HyPrColoc is designed for the colocalization of multiple complex traits at a locus, assuming a single shared causal variant, rather than the colocalization of multiple molecular phenotypes with a trait. The latter is more complicated owing to the diverse coverage of molecular phenotypes across the genome and multiple sites per molecular phenotype within a locus. Primo is another method of this kind that aims to identify SNPs associated with multiple traits.[15] To reduce spurious associations due to LD, Primo re-estimates the GWAS SNP effects conditioning on the top xQTLs in *cis*, leading to reduced power and computational efficiency. With the increasing availability of other types of xQTL data, such as chromatin accessibility QTL (caQTL) and alternative polyadenylation QTL (apaQTL),[16,22–27] methods that allow simultaneous integration of a larger number of molecular phenotypes with the GWAS data are essential to enhance the power to discover trait-associated molecular phenotypes and to infer the genetic mechanisms underlying complex traits.

In this study, we developed OPERA (Omics PlEiotRopic Association), an efficient and powerful approach that only requires summary statistics and reference LD to jointly analyze xQTL data from multiple omics layers with GWAS data to uncover plausible molecular mechanisms underlying GWAS loci. We show by simulations that OPERA yields well-controlled false discovery rate (FDR) and high power in detecting various patterns of pleiotropic associations. Here, pleiotropic association is defined as the association between a molecular phenotype and the trait through shared causal variant(s); hence, it includes the scenario where the molecular phenotype is a causative factor for the trait. We applied OPERA to summary-level xQTL data for seven molecular phenotypes and GWAS data for 50 complex traits and found that 50% of the GWAS signals were shared with at least one molecular phenotype, and all the analyzed molecular phenotypes were informative for interpreting the GWAS results.

RESULTS

Method overview

A full description of the OPERA method can be found in STAR Methods. In brief, our method combines multiple types of xQTL summary data (e.g., eQTL, mQTL, caQTL, hQTL, sQTL, apaQTL, and pQTL) with GWAS data to detect pleiotropic associations of molecular phenotypes with a complex trait at a GWAS locus, with the primary aim of providing mechanistic interpretation of the GWAS signal. We define a GWAS locus as a 2-Mb window centered at each of the independent GWAS signals identified by LD clumping[28] of GWAS summary statistics for a trait, with the overlapping loci merged together. OPERA is a Bayesian generalization of the SMR and HEIDI approach[8] to a multi-omics model, in which the molecular phenotypes are considered as multiple exposures and the complex trait is considered as the outcome. No assumption is made on the order of the molecular phenotypes regarding their effects on the outcome. For a locus with t molecular phenotypes, there are 2^t possible association patterns of the exposures with the outcome (referred to as "configurations," Figure 1). For example, in the OPERA analysis below with seven xQTL summary datasets, there were $2^7 (= 128)$ configurations in total. Assuming a point-normal mixture distribution for the effect of each exposure on the outcome (hereafter referred to as exposure effect), we model the likelihoods of all possible configurations based on the SMR estimates of the exposure effects and calculate the posterior probability for each configuration (PPC). PPCs are combined across configurations to compute a marginal posterior probability of association (PPA) for each exposure and a joint PPA for a combination of multiple exposures (integrating across the remaining exposures; STAR Methods). We focus primarily on PPA rather than PPC because PPC could be inflated for some configurations owing to a lack of power to detect small exposure effects (STAR Methods).

Computing PPA requires prior probabilities for the configurations (denoted by $\boldsymbol{\pi}$; see STAR Methods), which are unknown (i.e., we do not know the proportion of each configuration across the genome prior to the analysis). Given limited sample sizes of GWASs or xQTL studies, mis-specification of $\boldsymbol{\pi}$ can lead to false positives in the posterior inference.[14] In theory, we can perform a full Bayesian analysis to estimate all the model parameters by Markov chain Monte Carlo. In practice, however, such a full Bayesian analysis is extremely challenging to achieve because of the heavy computational burden owing to the presence of a large number of sites at a locus for some exposures, such as DNAm probes or chromatin accessibility peaks, and the complex correlation structure between sites within and, potentially, across loci (Data S1). Therefore, we propose a two-stage strategy to approximate the full Bayesian inference. In the first stage, we run a full Bayesian model to estimate $\boldsymbol{\pi}$ using a set of quasi-independent loci (i.e., non-overlapping loci with the presence of all molecular phenotypes in each locus) across the genome. In the second stage, we compute the marginal PPA for each molecular phenotype and the joint PPA for multiple molecular phenotypes based on the estimated $\boldsymbol{\pi}$ values from stage 1. For a molecular phenotype with multiple sites at a locus, a randomly selected site is used in the stage-1 analysis, and all the sites are then analyzed one by one, in combination with the other molecular phenotypes, in the stage-2 analysis (STAR Methods). We report the average marginal and joint PPA when a site or combination of sites is involved in multiple tests.

Like the single-exposure association identified by SMR that can be driven by LD between distinct causal variants for the exposure and outcome, the joint-exposure association identified by OPERA can also be due to LD between xQTL and GWAS SNPs or between xQTL SNPs across exposures (Figure S1A). To filter out such associations, we perform a multi-exposure HEIDI test, extended from the single-exposure HEIDI test,[8] on the OPERA associations with PPA > 0.9 (STAR Methods). The null hypothesis for the multi-exposure HEIDI

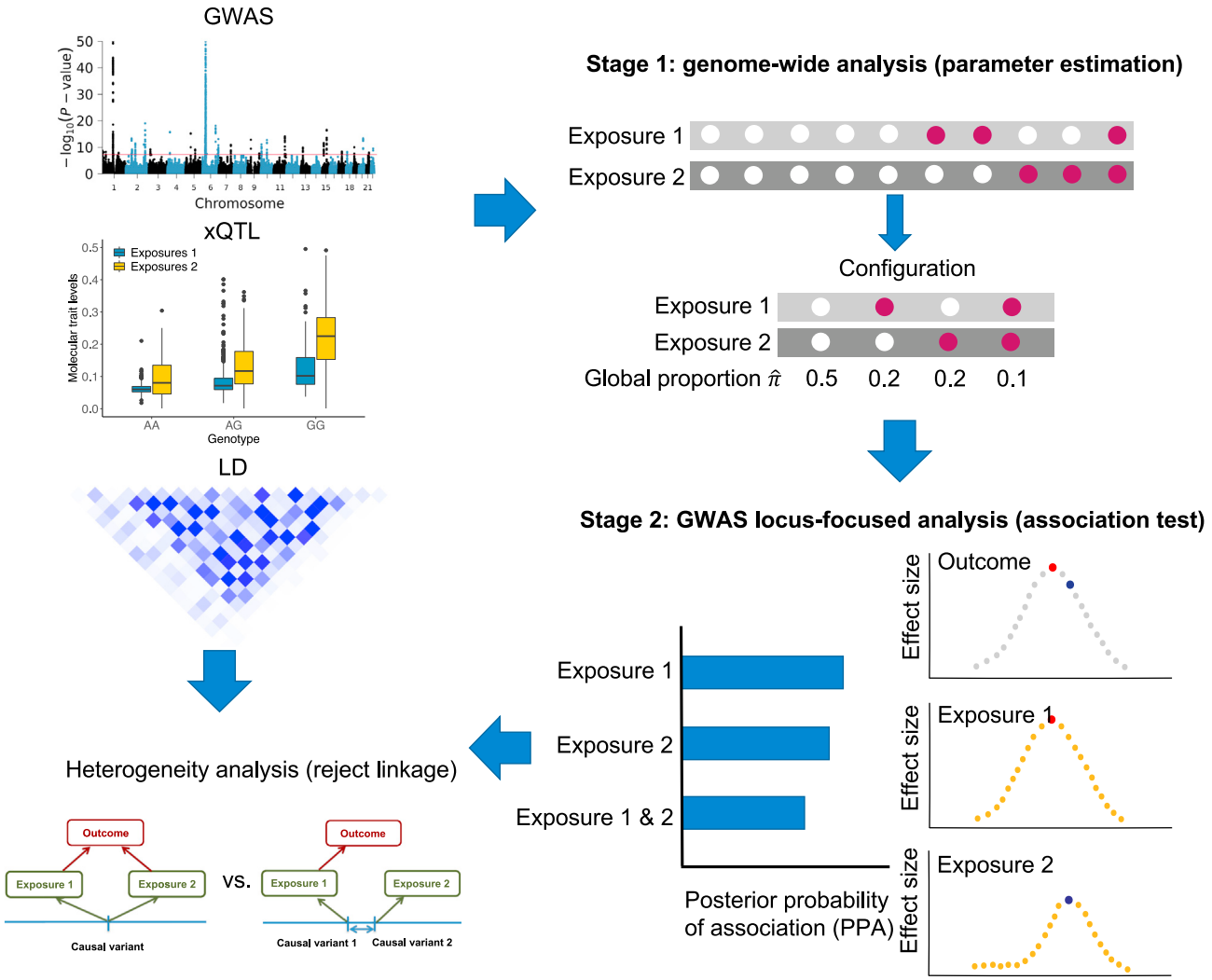


Figure 1. Schematic overview of OPERA

OPERA combines GWAS summary statistics, multiple xQTL summary statistics from multiple omics layers, and reference LD to identify molecular phenotypes and their combinations that are associated with a complex trait of interest because of pleiotropy. The OPERA analysis consists of three steps. OPERA first estimates the global proportions of possible configurations (i.e., $\boldsymbol{\pi}$) using a set of quasi-independent loci across the genome. Using the estimated $\boldsymbol{\pi}$ and data likelihood under each configuration, OPERA then computes the posterior probability for each configuration (PPC_γ) for all possible combinations between molecular phenotypes at a GWAS locus. Thereafter, PPCs across configurations are combined to compute the marginal posterior probability of association (PPA) for each exposure and the joint PPA for multiple exposures (STAR Methods). For associations with high PPAs (e.g., PPA > 0.9), OPERA performs the heterogeneity test (multi-exposure HEIDI test) to reject associations that are caused by linkage. The marginal and joint molecular phenotype associations with high PPAs and passed multi-exposure HEIDI test are accepted as pleiotropic/causal associations. The solid symbols represent the exposure sites associated with the outcome, and the hollow symbols indicate that the exposure sites have a null effect on the outcome.

test is that the association between exposure(s) and outcome is due to pleiotropy at the same causal variant(s) (referred to as a homogeneous causal model; FiguresS1B and S1C). To claim pleiotropic associations under a homogeneous causal model, we only accept OPERA associations that pass the multi-exposure HEIDI test (e.g., $p_{HEIDI} > 0.01$; justification for using this threshold is discussed in Wu et al.[16]). Like the single-exposure HEIDI test, the multi-exposure HEIDI test is generally conservative, meaning that it can reject pleiotropic models when the GWAS and xQTL causal variants are partially shared (Figure S1D), but its power to detect heterogeneity decreases with the increase of LD between the GWAS and xQTL causal variants.[8] To mitigate the cross-exposure leakage effect, we used the joint-exposure effects (i.e., the joint SMR effects), which can be estimated from xQTL and GWAS summary statistics employing reference LD (STAR Methods, Data S2, and Figures S2A and S2B), for estimating $\boldsymbol{\pi}$ in the stage-1 analysis and detecting exposure-outcome associations in the stage-2 analysis. For the associations detected in the stage-2 analysis, we further applied the multiple-exposure HEIDI test to reject those inconsistent with a homogeneous causal model. We have implemented OPERA in a user-friendly software tool

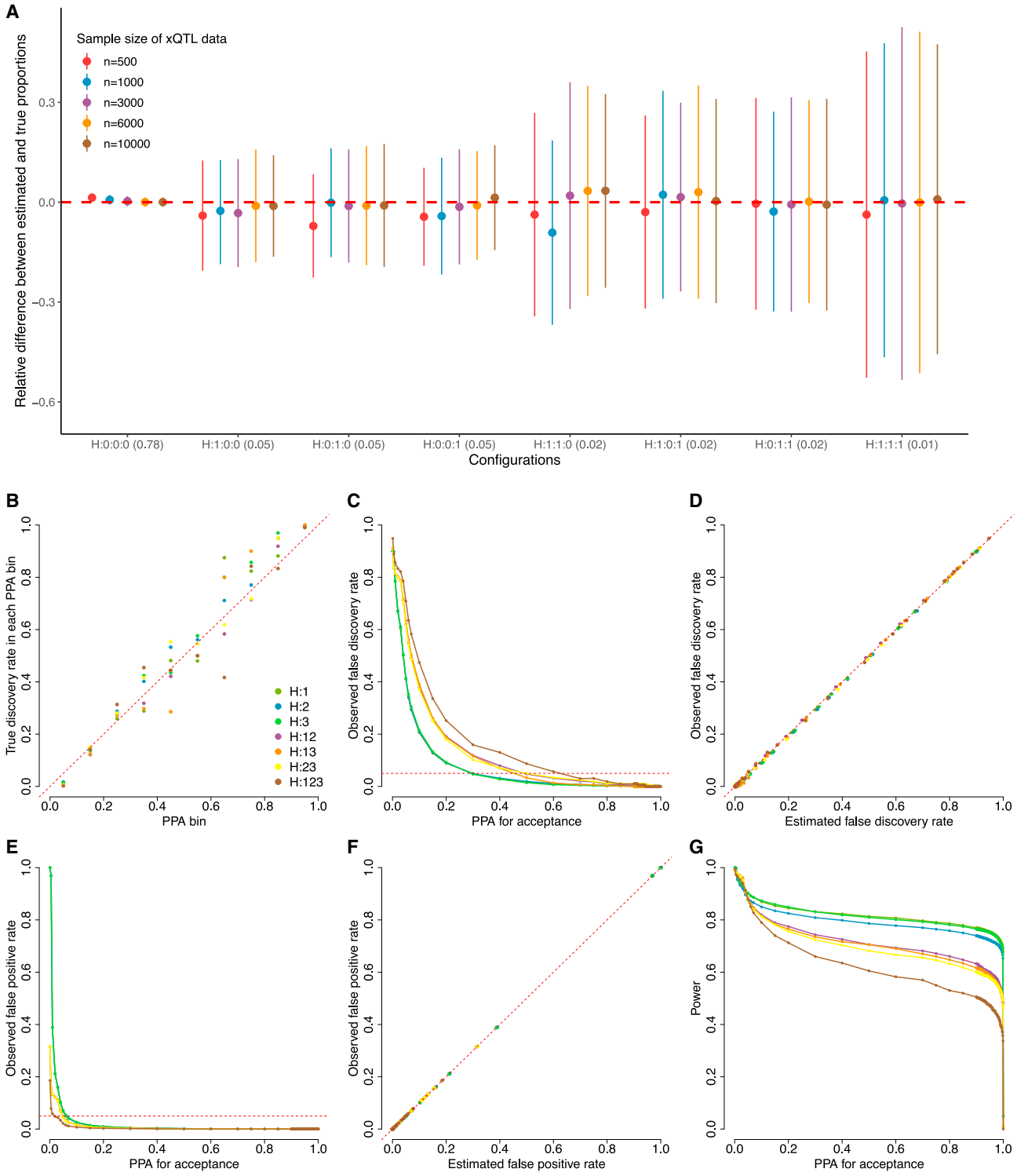


Figure 2. Performance of OPERA in estimating the proportion of loci under each configuration and detecting multi-omics associations

(A) Estimation of π from OPERA using simulations based on the imputed genotype data from the UK Biobank, where π is the vector of proportions attributed to each configuration. Shown are the results based on 400 simulated independent loci, with three molecular phenotypes at each locus. Each dot shows the mean difference between the estimated and true proportions across 100 simulation replicates, and each solid line represents ±1 SD. The x axis shows the configurations with their true proportions in parentheses. For example, the proportion of configuration where none of the three exposures are associated with complex

(legend continued on next page)

available at https://github.com/wuyangf7/OPERA. The whole analysis only requires xQTL and GWAS summary statistics and reference LD.

Assessing the performance of OPERA in estimating prior probabilities

To assess the performance of OPERA in estimating the proportion of loci under each configuration (output from the stage-1 analysis), we started by performing simulations based on unlinked variants. We simulated 400 independent loci and three molecular phenotypes at each locus (one exposure site per molecular phenotype and one xQTL per site). A complex trait (i.e., outcome) was simulated by assuming an additive model for the exposure effects across loci (STAR Methods). The true association statuses of the exposures at each locus were sampled at random given a setting of $\boldsymbol{\pi}$, and the exposure effects on the outcome were sampled from a point-normal mixture distribution (STAR Methods). We set 0.78, 0.05, 0.02, or 0.01 as the proportion of loci (i.e., the true values of $\boldsymbol{\pi}$) with 0, 1, 2, or 3 exposures associated with the outcome, respectively. The results showed that $\widehat{\boldsymbol{\pi}}$ values were highly accurate (Figure S3) and not significantly different from the true values of $\boldsymbol{\pi}$ (smallest p = 0.31), with the accuracy measured by the relative difference between $\widehat{\boldsymbol{\pi}}$ and $\boldsymbol{\pi}$ ($\Delta_r = (\widehat{\boldsymbol{\pi}} - \boldsymbol{\pi})/\boldsymbol{\pi}$). We focused on the relative difference because it accounts for the uneven distribution of $\boldsymbol{\pi}$ (see Figure S4 for the absolute difference between $\widehat{\boldsymbol{\pi}}$ and $\boldsymbol{\pi}$, $\Delta_a = \widehat{\boldsymbol{\pi}} - \boldsymbol{\pi}$). However, when the sample size is very small (e.g., n = 300), we observed a slight deflation for the alternative hypotheses, and the deflation is inversely proportional to the number of associated exposures due to a lack of power (Data S3). We also performed simulations with sample sizes and xQTL effect sizes observed from real xQTL data for five molecular phenotypes and observed small and non-significant Δ_r (Figure S5). We further tested our method in a more extreme case where half of the true $\boldsymbol{\pi}$ values were zero in the simulation and still found the $\widehat{\boldsymbol{\pi}}$ value to be robust (Figure S6).

We next assessed our method in estimating $\boldsymbol{\pi}$ using simulations based on the imputed genotype data from the UK Biobank.[29] Unlike the simulation based on unlinked variants where there was only one variant per exposure at each locus, here we simulated a locus by picking a 500 kb genomic region at random and sampled the number of causal variants from a Poisson distribution (mean = 2) for each of the three exposures, allowing LD between causal variants within and across exposures (STAR Methods). We first simulated three well-powered xQTL datasets with an even sample size (n_{xQTL} = 3,000) and compared $\widehat{\boldsymbol{\pi}}$ estimated using SMR (without HEIDI) with a p_{SMR} threshold of 0.05/400, using OPERA with marginal SMR effects and using OPERA with joint SMR effects. The results showed that $\widehat{\boldsymbol{\pi}}$ estimated using OPERA based on the joint SMR effects, which accounts for cross-exposure leakage effects (STAR Methods), were unbiased, whereas the estimates using SMR were biased (Figure S7), which is expected because of selection bias. We further simulated xQTL datasets with a wide range of sample sizes and observed results consistent with those from the simulations based on unlinked variants (Figure 2A).

Assessing the performance of OPERA in detecting multi-omics pleiotropic associations

We have shown above that the estimation of $\boldsymbol{\pi}$ from the stage-1 analysis of OPERA was sufficiently accurate. Here we evaluated the FDR, false positive rate (FPR), and power of the stage-2 analysis of OPERA in detecting associations of molecular phenotypes with a trait, using $\widehat{\boldsymbol{\pi}}$ from the stage-1 analysis (STAR Methods). It has been shown previously that Bayesian posterior probabilities such as PPA can be used to control FDR in a multiple-test setting.[30–32] We first evaluated the relationship between PPA and true discovery rate (TDR) in ten PPA bins, where TDR in each PPA bin is computed as the number of causal exposures divided by overall number of exposures in the bin. We observed good agreement between PPA and TDR (r = 0.98, Figure 2B). Next, we estimated the FDR at a given PPA threshold using an empirical Bayesian approach[15,30,32] (STAR Methods), and compared it with the observed FDR (Figure 2C), i.e., the proportion of false positives among the identified exposures at a PPA threshold. Figure 2D shows a strong concordance between the estimated and observed FDR, meaning that in practice, a PPA threshold can be analytically solved once a desired FDR level is determined. Furthermore, we proposed an approach to estimate FPR at a PPA threshold, given the estimated proportion of the null model from the stage-1 analysis (STAR Methods). Our estimated FPR was in almost perfect agreement with the observed FPR (Figure 2E), i.e., the proportion of false positives under the null (Figure 2F), even at a relatively high significance level of 5×10^{-5} (Figure S8). This result is appealing to researchers who prefer controlling FPR (e.g., using a genome-wide significance level) over FDR. In this study, we used a PPA threshold of 0.9, which sufficiently controlled the FDR below 0.05 and gave a power of at least 0.5 for detecting the simulated causal exposures (Figure 2G), where the mean exposure effect variance estimated from the real data was used as the true value in the simulation. In practice, a different PPA threshold can be selected depending on the desired FDR or FPR level for the study.

One major advantage of the joint analysis of multi-omics xQTL data is that it can substantially increase power. To demonstrate this, we compared the power of the joint analysis with that of

trait, i.e., H:0:0:0, is 0.78. The red dashed line represents zero difference between the estimated and true proportions. The color of the box denotes the sample size of xQTL studies for the three exposures.

(B) Relationship between PPA and the true discovery rate across ten PPA bins. Colors show the results of different association hypotheses, e.g., H:1 indicates the marginal association of the first exposure with the outcome, and H:123 indicates joint association of three exposures with the outcome. The red line represents y = x.

(C, E, and G) The observed false discovery rate (FDR), false positive rate (FPR), and power along with an increasing PPA threshold. In (C) and (E), the red dashed line represents an FDR or FPR of 0.05.

(D and F) Consistency of the estimated and observed FDR (or FPR) from OPERA given a PPA threshold, with the red line being y = x.

Shown in (B) to (G) are the results computed from 100 simulations using a sample size of 300,000 for GWASs and 1,000 for the xQTL studies.

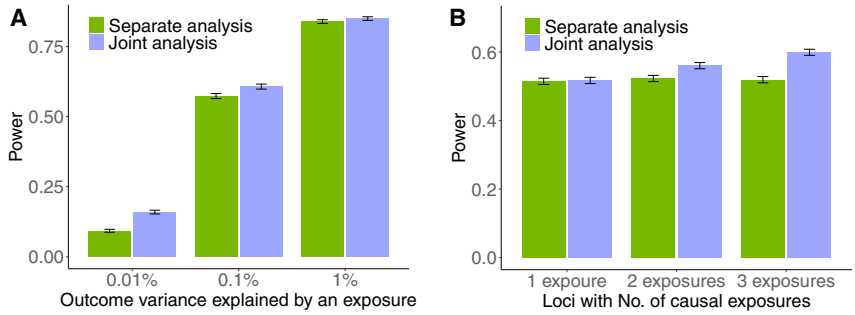


Figure 3. Increased discovery power using joint analysis of multiple xQTL datasets

(A) The statistical power of OPERA for detecting causal exposures in simulations using all xQTL data jointly or each xQTL dataset separately. The error bar is the estimated standard error of the mean across simulation replicates.

(B) The power of OPERA in simulations with different numbers of causal exposures.

separate OPERA analysis of each molecular phenotype individually. We opted for the separate OPERA analysis instead of SMR as it allows for comparison based on the same PPA threshold, and because the separate OPERA analysis showed receiver-operating characteristic (ROC) curves very similar to those of SMR (Figure S9). In the simulations, we set different values for the proportion of variance of the trait explained by each exposure measure on average (R^2_{xy}) for the three molecular phenotypes (STAR Methods). The results showed that compared with the separate analysis, the joint analysis consistently increased the discovery power for each molecular phenotype, particularly for the molecular phenotype with the smallest R^2_{xy} (power increased by 1.7-fold; Figure 3A). We further compared the power of joint and separate analyses in the simulations with varying numbers of causal exposures. Our results showed a significant increase in power with an increasing number of causal exposures for the joint analysis but not for the separate analysis (Figure 3B).

We next calibrated the multi-exposure HEIDI test by simulation. Under the null model where the causal exposures and the outcome shared the same causal variant(s) (homogeneous causal model; Figures S1B and S1C), the multi-exposure HEIDI test statistics were not inflated regardless of whether we tested a single exposure separately or two exposures jointly (Figures S10A and S10B). We then simulated the alternative model in which only the first exposure had a causal effect on the outcome (heterogeneous causal model), and the causal variant of the second exposure was in LD ($0.1 < r^2 < 0.9$) with that of the first exposure (Figure S1A). When each exposure was tested separately, the multi-exposure HEIDI test statistics were not inflated for the first exposure (Figure S10C) but inflated when two exposures were tested jointly (Figure S10D), indicating that the multi-exposure HEIDI test correctly accepted the configuration that only the first exposure was in pleiotropic association with the outcome and rejected the configuration that the two exposures were jointly in pleiotropic association with the outcome.

To assess the robustness of OPERA, we performed extensive simulations that involved various scenarios, such as reducing the xQTL sample size, utilizing sample size and effect size as observed from real data, treating 10% of exposure sites as missing, simulating exposures with multiple sites at a locus, using a different significance threshold to estimate the prior variance, and simulating correlated effects between exposures. The simulation results showed that OPERA remains robust in all scenarios (Data S4 and Figures S11–S16). Furthermore, we examined the potential impact of cellular heterogeneity (i.e., cell-type-specific xQTL or exposure effect) on the multi-exposure HEIDI test and found that the multi-exposure HEIDI test remained robust in distinguishing the homogeneous and heterogeneous causal models even in the presence of cellular heterogeneity (Data S5 and Figure S17).

Comparison of OPERA with other colocalization methods

We compared OPERA with three state-of-the-art colocalization methods, i.e., MOLOC,[20] HyPrColoc,[21] and Primo,[15] all of which report PPC for a substantial number of configurations at SNP level to make a distinction between homogeneous and heterogeneous causal models. For a fair comparison, we collapsed the PPC of MOLOC, HyPrColoc, and Primo into PPA, following the same procedure as in OPERA (STAR Methods). We used the same summary statistics as input for all methods and applied the default prior probabilities for MOLOC (version 0.1.0) and HyPrColoc (version 1.0), while Primo (version 0.2.1) estimated its prior probabilities. We provided Primo the same LD reference used in the OPERA analysis for conditional association analysis. The results showed that OPERA outperformed MOLOC, HyPrColoc, and Primo, as evidenced by the higher area under the ROC curve and higher power at the same levels of FPR (Figure 4). Moreover, OPERA is computationally much more efficient, being approximately 6 times faster than HyPrColoc, 29 times faster than Primo, and 56 times faster than MOLOC in analyses involving three molecular phenotypes (Table S1). While MOLOC and Primo exhibited a substantial increase in computing time when the number of exposures increased from three to four, OPERA's computing time only increased slightly (Table S1), indicating that the rate of increase was slower for OPERA compared with MOLOC and Primo.

Identifying multi-omics pleiotropic associations from real data

We applied OPERA to test for pleiotropic associations of seven molecular phenotypes with 50 complex traits. The GWAS summary data were from the largest available GWAS meta-analyses[29,33–43] for 19 complex traits and from the UK Biobank for 31 complex traits (Table S2). The blood eQTL summary data were from the eQTLGen study[22] (n = 31,684) and the CAGE study[44] (n = 2,765), and the caQTL summary data from

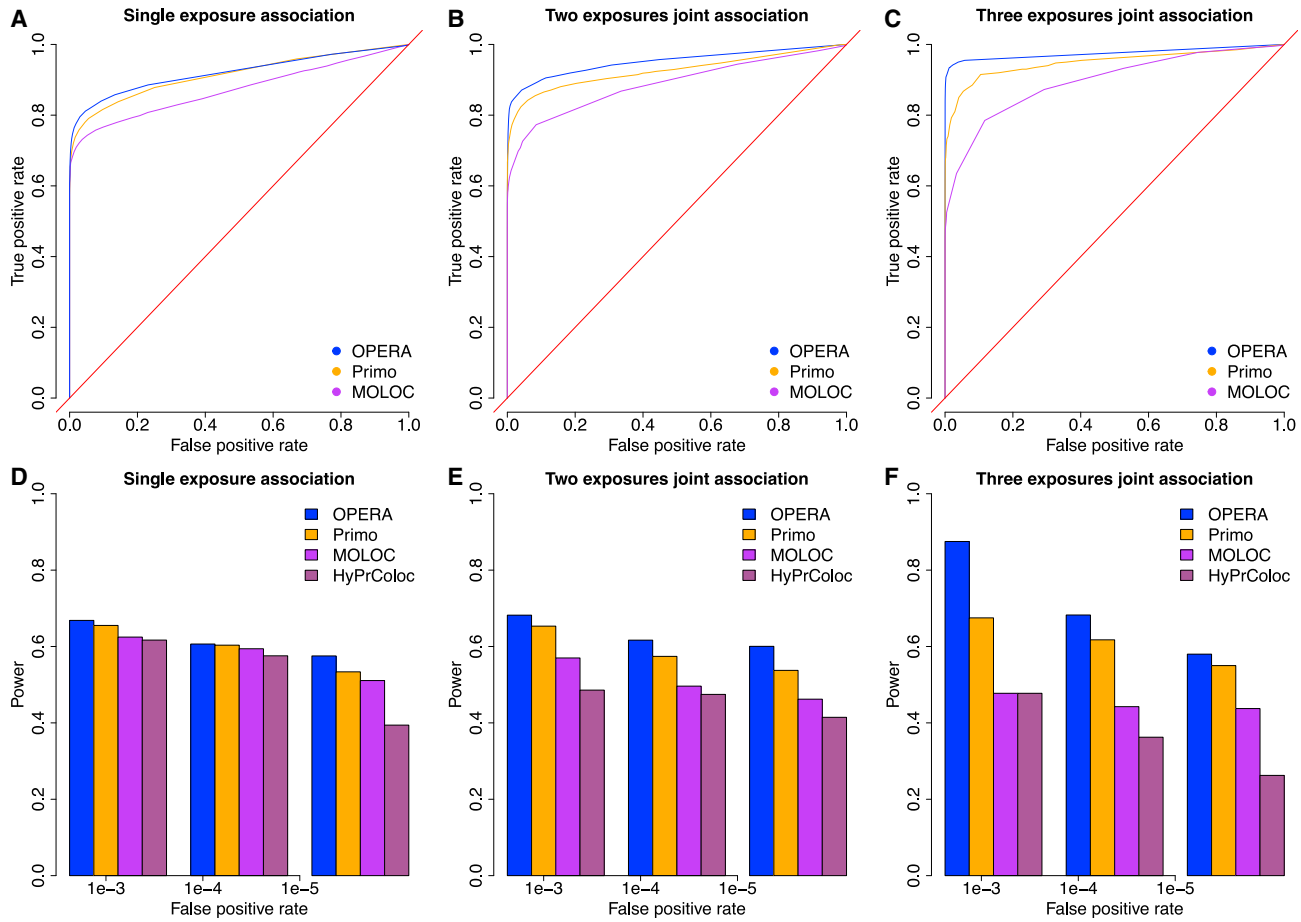


Figure 4. Comparison of OPERA with three existing methods: MOLOC, Primo, and HyPrColoc

(A–C) The x axis represents the false positive rate, and the y axis represents the statistical power (as measured by true positive rate) of the methods.

(D–F) The power of the methods (y axis) at different levels of false positive rate (x axis). We cannot compare OPERA with HyPrColoc with a full ROC curve because HyPrColoc only reports significant PPA that passes the internal selection criterion.

lymphoblastoid cell lines were from Kumasaka et al.[24] (n = 100). The mQTL summary data were generated from a meta-analysis of two blood mQTL datasets from McRae et al. (n = 1,980).[16,26] Both sQTL and apaQTL[27] summary data were from the whole blood samples included in the GTEx project v8[45] (n = 670). The summary-level hQTL (including H3K27ac and H3K4me1) data for naive CD14^{+} monocytes were from the BLUEPRINT project[25] (n = 200), and the plasma pQTL summary statistics were from the INTERVAL study[23] (n = 3,301). All the xQTL effects were in standard deviation (SD) units. After applying quality control filters (STAR Methods), we retained 15,511 genes for eQTLGen, 9,538 genes for CAGE, 13,834 chromatin accessibility peaks, 90,827 DNAm sites, 6,638 RNA splicing events, 1,050 3′ UTR alternative polyadenylations, 18,153 histone marks, and 687 proteins with at least one xQTL with $p < 5 \times 10^{-8}$. To compute the joint SMR effects, we performed GCTA-COJO[46] analysis to identify quasi-independent xQTL signals for each molecular phenotype. The proportion of variance in a molecular phenotype explained by an xQTL COJO signal, on average, increased in the order of pQTL, eQTL (CAGE, n = 2,765), mQTL, apaQTL, sQTL, hQTL, and caQTL, while the number of xQTL COJO signals decreased in reverse order (Figures S18A and S18B), an order consistent with a higher level of complexity in genetic architecture for molecular phenotypes functionally more distal from DNA. However, it should be noted that sample size also contributed to the observed differences. When the eQTLGen dataset was used instead of CAGE, the eQTL result appeared to be an outlier (Figures S18C and S18D), likely due to the much larger sample size of the eQTLGen data than that of the other xQTL datasets, leading to a discovery of many eQTLs with smaller effects. We used the eQTLGen data for discovery because its sample size is much larger than that of the CAGE data.

In the stage-1 analysis, the global proportions of the 128 configurations for each complex trait (STAR Methods) need to be estimated from genome-wide quasi-independent loci, which requires the presence of all seven molecular phenotypes at each locus. Thus, we started with the pQTL data because of their low coverage (698 proteins in total), defined a 500-kb region/locus centered around the coding gene of each protein, and excluded loci at which any of the other molecular phenotypes

was missing. For the molecular phenotypes with multiple sites, we randomly sampled one site at each locus. On average, across the 50 complex traits, the proportion of loci with at least one molecular phenotype associated with a complex trait (π_{alt} = 1 − [proportions of loci with no association]) was 21.0% (Figure S19). Note that π_{alt} should be interpreted as the proportion of xQTLs associated with complex traits rather than the proportion of GWAS loci associated with molecular phenotypes because not all the selected loci harbored GWAS signals. The estimate of π_{alt} varied substantially across the complex traits, ranging from 8.9% (SE = 2%) for Alzheimer's disease to 61.9% (SE = 4%) for height. We then investigated whether $\widehat{\pi}_{alt}$ varied with GWAS sample size. We observed a significant decrease of $\widehat{\pi}_{alt}$ when analyzing GWAS data for height and body mass index (BMI) with smaller sample sizes[47,48] (Figure S20A). Given the limited number of quasi-independent loci with all seven molecular phenotypes present (ranging from 148 to 168 across complex traits), we examined the robustness of the estimation by comparing $\widehat{\pi}_{alt}$ with the estimate from the analysis of two exposures, for which more independent loci were available (ranging from 400 to 134,103, depending on the molecular phenotype combination). To compare the results between the seven-exposure and two-exposure analyses, we computed a marginal $\widehat{\pi}_{alt}$ for each molecular phenotype by summing up the relevant elements in $\widehat{\pi}$. The result showed that the marginal-proportion $\widehat{\pi}_{alt}$ for each molecular phenotype estimated from the seven-exposure analysis was highly consistent with that estimated from the two-exposure analysis (Figure S20B).

We then used $\widehat{\pi}$ from the stage-1 analysis to perform the stage-2 analysis, followed by the multi-exposure HEIDI test, at the independent GWAS loci after LD clumping (LD r^2 threshold = 0.01 and window size = 1 Mb) for the 50 complex traits. For ease of computation, we included only the molecular phenotype sites that passed the nominal SMR threshold (i.e., $p_{SMR} < 0.05$) within 1 Mb distance of each GWAS signal in either direction, resulting in a total number of ~1.1 million pleiotropic association tests. We limited our analysis in such a 2 Mb window because for some molecular phenotypes, the xQTL data were only available for SNPs within 1 Mb of the target site. At a PPA threshold of 0.9 and p_{HEIDI} threshold of 0.01, we identified 42,305, 61,794, 218,809, 215,344, 112,052, or 23,782 pleiotropic associations for which 1, 2, 3, 4, 5, or 6 molecular phenotypes were associated with a complex trait, respectively, but none of the GWAS signals was explained by all seven molecular phenotypes jointly (Figure 5A). The estimated FDR for each marginal or joint association was <0.05, with an average of 0.023 across association combinations and traits, meaning that the overall FDR of the whole experiment was also <0.05. Across 50 complex traits, we observed a strong correlation between the number of associated molecular phenotypes and the number of independent GWAS loci (Figure S21). The number of identified pleiotropic associations (42,305) was larger than the number of GWAS loci examined (12,068), suggesting, on average, ~3.5 trait-associated molecular phenotypes per GWAS locus. These pleiotropic associations consisted of 17,619 unique molecular phenotype measures (Tables S3–S9), of which 7,901 were not detected by a separate analysis of each omics layer individually. For example, the joint analysis with the seven layers of xQTL data identified 2,169 unique genes (Table S3), 793 of which were not detected by the analysis using eQTL data only, demonstrating the increased power of gene discovery through the joint analysis of multi-omics xQTL data. For each of the other six types of xQTLs, the seven-exposure analysis consistently discovered more trait-associated sites than the single-exposure analyses (Figure S22), in line with the simulation results, particularly for exposures with small effects (Figure 3A). Of the 2,169 prioritized genes, 1,493 genes (68.8%) were associated with at least one *cis*-regulatory molecular phenotype (i.e., DNA methylation, histone modification, or chromatin accessibility), indicating that the trait-associated genes were largely modulated by *cis*-regulatory elements. We identified 57 proteins that showed pleiotropic association with the complex traits (Table S4), 26 (45.6%) of which were detected in joint pleiotropic association with a *cis* gene, similar to the reported overlapping proportions between *cis*-pQTLs and *cis*-eQTLs from previous studies.[23,49]

The pleiotropic associations we identified included genes and proteins that were reported previously for the corresponding traits. For example, 32 of the 72 schizophrenia (SCZ)-associated genes identified by eQTLs using the seven-exposure analysis overlapped with the genes reported in a recent study by integrating a newer release of the SCZ GWAS data with the eQTLGen data using SMR, including *SNX19* and *RERE* that have been validated by other studies using different methods.[4,20] We identified protein CD33.3166.92.1 for Alzheimer's disease using the plasma pQTL data, consistent with the function of the *CD33* gene, which is to increase the number of activated human microglia.[50] Some proteins that are the targets of approved drugs were found to be associated with the corresponding disease, such as interleukin-6 receptor for coronary heart disease.[51] We detected several genes and proteins that showed pleiotropic effects on multiple traits, e.g., MAPK3.2855.49.2 protein, which was identified to be associated with three traits, including red blood cell counts, height, and BMI, implying a shared molecular mechanism between the traits at the *MAPK3* locus. We also performed four sets of validation analyses to replicate or validate the pleiotropic associations detected by the analyses above (Data S4).

Understanding the molecular mechanisms at GWAS loci

The detection of pleiotropic associations of different combinations of molecular phenotypes with complex traits provides new insights into the molecular mechanisms underlying the GWAS signals. Across the 50 complex traits, we observed 50.0% of GWAS loci had at least one molecular phenotype association in the joint analysis, with the proportion varying from 22.2% for hemorrhoids to 63.3% for waist circumference adjusted for BMI (Figure 5B). Notably, 24.9% of GWAS loci had at least two molecular phenotypes jointly associated with a trait, suggesting the synergistic actions of molecular phenotypes in mediating the genetic effects on complex traits. Only 19.8% of GWAS loci had a transcriptional abundance association in blood (Figure 5B), likely because of spatiotemporal effects or mechanisms beyond the genetic control of transcriptional abundance, in line with the conclusion from a prior study of autoimmune diseases.[52] About 39.0% of GWAS loci had a DNA

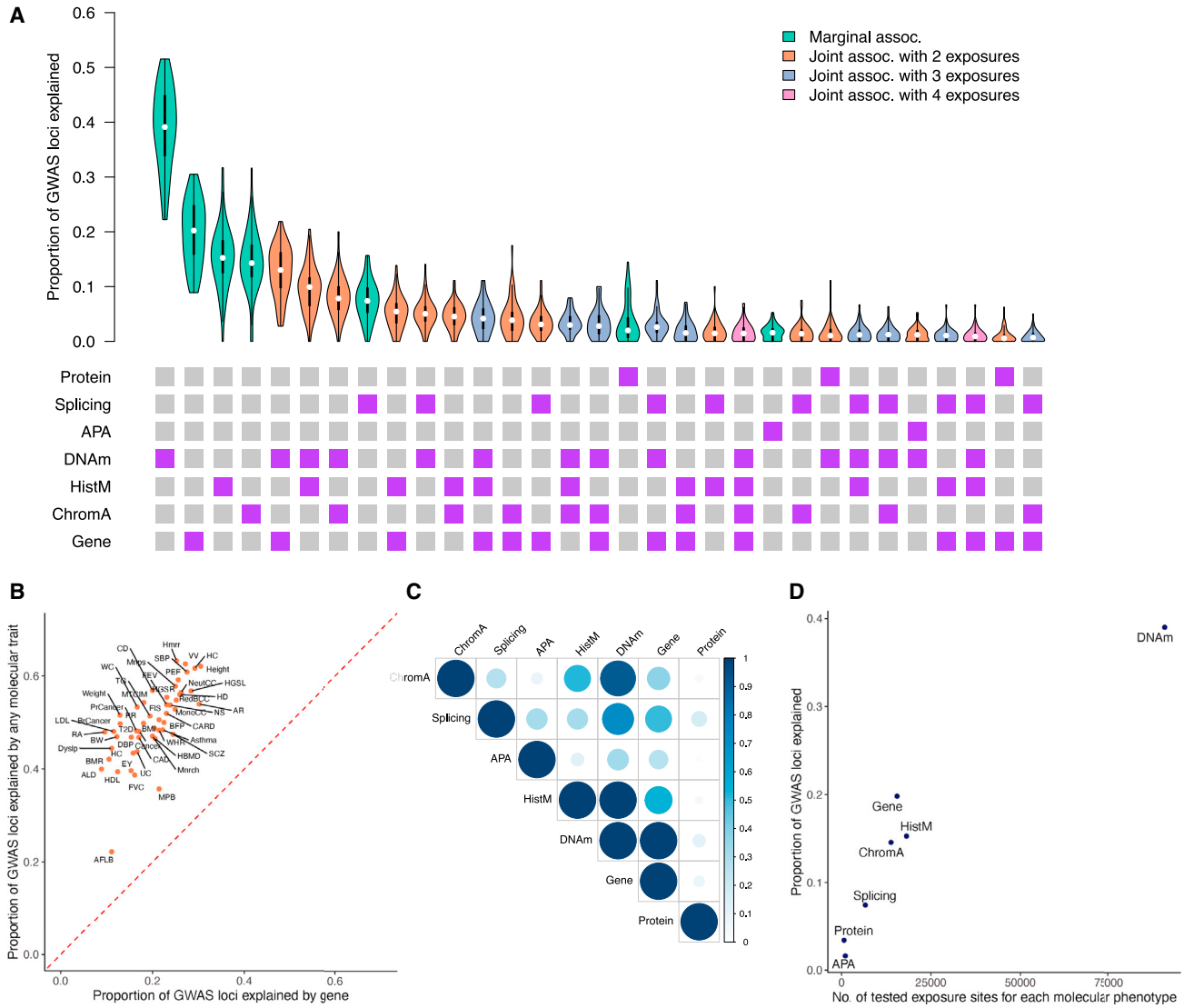


Figure 5. Proportion of GWAS loci explained by the detected pleiotropic associations in the joint analysis with seven molecular phenotypes

(A) Proportion of GWAS loci explained by each combinatorial pleiotropic association. The quantified proportions are based on results of passed PPA and multi-exposure HEIDI rather than PPC, thus these proportions are not independent and the sum of them is not equal to 1. The x axis shows the association hypotheses with different molecular phenotype combinations; only 32 combinations with largest proportion of GWAS loci explained are shown in the plot. The purple (gray) block represents the presence (absence) of each molecular phenotype, including protein, RNA splicing (Splicing), 3′ UTR alternative polyadenylation (APA), DNA methylation (DNAm), histone modification (HistM), chromatin accessibility (ChromA), and gene expression (Gene). The y axis shows the proportion of GWAS independent loci explained by each marginal or joint association hypothesis across 50 complex traits. In each violin box, the center line shows the median, box limits are the upper and lower quartiles, whiskers represent 1.5× interquartile range, and individual points are outliers.

(B) Relationship between proportion of GWAS loci explained by each type of molecular phenotype and number of tested exposure sites for each molecular phenotype.

(C) Proportion of GWAS loci explained by any molecular phenotype or gene only for the analyzed 50 complex traits. The estimated proportions are dependent on the sample sizes of xQTL studies in two ways. First, the estimated proportion for each marginal or joint association is expected to be a lower bound given the limited xQTL sample sizes for most molecular phenotypes. Second, the relative differences between molecular phenotypes partially reflect the differences in power between xQTL studies in addition to the differences between the true proportions. The acronyms for complex traits can be found in STAR Methods.

(D) Proportion of overlap of molecular phenotypes in the identified joint associations. The proportion is calculated as the number of identified joint associations for each pair of molecular phenotypes divided by the square root of the product of the number of identified marginal associations for each molecular phenotype. Similar to (A), the estimated overlapping proportions are likely affected by sample sizes and coverage of the molecular phenotypes.

methylation association, and DNA methylations were actively involved in the joint associations with other molecular phenotypes (Figure 5C), although the sample size of the mQTL data (n = 1,980) was much smaller than that of the eQTL data (n = 31,684), highlighting the importance of integrating mQTL data to dissect molecular mechanisms at GWAS loci. Among

the seven xQTL datasets analyzed, apaQTL, pQTL and sQTL accounted for the three smallest proportions of the GWAS loci (1.6%, 3.4%, and 7.5%, respectively), mainly due to the limited sample size (n = 670 for both apaQTLs and sQTLs) and limited number of exposure sites included in the analysis (n = 1,050, 6,638, and 687 for apaQTLs, sQTLs, and pQTLs, respectively; Figure 5D). Despite the small sample sizes of the hQTL (n = 200) and caQTL data (n = 100), 15.3% and 14.6% of the GWAS signals were shared with histone modification and chromatin accessibility sites, respectively. Overall, all the analyzed molecular phenotypes played a role in explaining the GWAS signals. The remaining unexplained GWAS loci (~50%) open opportunities for future work to fully characterize the molecular mechanisms underlying GWAS loci through increasing the diversity and sample sizes of xQTL data (see below for more discussion).

We also observed association signals of multiple molecular phenotypes shared with a single GWAS signal. The *MANBA* gene for type 2 diabetes (T2D) was such an example. We showed that *MANBA* was associated with T2D jointly with five other types of molecular phenotypes, except for alternative polyadenylation (joint PPA for all six molecular phenotypes = 0.998) (Figure S23). The SNP-association pattern from GWAS was consistent with that from all the six xQTL studies, implying a plausible regulatory mechanism through which the SNP effect on trait is mediated. A similar example was found for SCZ at the *RERE* locus (Figure S24). There was another example in which five types of molecular phenotypes were jointly associated with prostate cancer at the *MSMB* locus (joint PPA = 0.952), where the SNP-association signals were consistent across the GWAS and all the xQTL data (Figure 6). Of note, the association between the MSMB.10620.21.3 protein and prostate cancer was identified using the blood pQTL data, but there was no significant eQTL for *MSMB* in the eQTLGen data (blood tissue), consistent with the observation that the expression level of *MSMB* was prostate-tissue specific (Figure S25). Adding the GTEx[53] prostate eQTL data for a six-exposure OPERA analysis identified significant pleiotropic association of the expression level of *MSMB*, among other molecular phenotypes (i.e., chromatin accessibility, histone modifications, DNA methylations, protein abundance), with prostate cancer. The estimated effects of both the protein and gene on prostate cancer were protective. We further checked the expression of *MSMB* in The Cancer Genome Atlas database and observed that the expression of *MSMB* in the normal prostate tissue was almost 2-fold higher than that in tumor tissue (Figure S25), consistent with the protective role of *MSMB* against prostate cancer.

DISCUSSION

In the present study, we have developed a powerful and efficient method for joint analysis of summary statistics from GWASs and multi-omics xQTL studies to detect pleiotropic associations between molecular intermediates and a complex trait of interest, with the primary aim of providing mechanistic interpretations of GWAS signals. Compared with the separate analyses of each layer of the xQTL data analyzed with the GWAS data one by one, the joint analysis of multi-omics xQTL data gains power, enabling the detection of molecular phenotypes with relatively weak effects. Compared with other existing integrative methods that can analyze multiple xQTL datasets, our method features the ability to estimate the global proportions of omics-trait association configurations from the data, control spurious associations due to the cross-exposure leakage effect, and relax the strict assumption about the number of causal variants per locus (e.g., MOLOC assumes at most one causal variant for each phenotype at a GWAS locus). Since OPERA uses the SMR estimates of exposure effects as the input, estimates from other MR approaches can also be used as long as they can provide the joint estimates of multiple exposures from summary statistics. Unlike the methods that assess the concordance of association signals of many SNPs at a locus among phenotypes, e.g., MOLOC and Primo,[15] OPERA is established on the SMR estimates of the exposure effects, which are at a higher level than the SNP effects, allowing for efficient integration of a number of molecular phenotypes simultaneously with GWAS data. Furthermore, the multi-exposure HEIDI test allows us to select high-PPA associations with evidence of pleiotropy (association between exposures and outcome due to shared causal variant [s]). Although the multi-exposure HEIDI test can indicate whether a GWAS signal shares the same underlying causal variant(s) with other molecular phenotypes, it cannot further prioritize causal variant(s). Methods that integrate functional annotation data with GWASs (e.g., latent probit model[54]) can be used to fine-map the putative causal variants at the loci detected by OPERA.

Through a large-scale analysis with the largest available summary-level GWAS data for 50 complex traits (n = 20,833–766,345) and xQTL data from seven omics layers (n = 100–31,684), we identified pleiotropic associations of 17,619 molecular phenotype measures with 50 complex traits, with an important observation that approximately 50% of the GWAS signals were shared with at least one molecular phenotype. The estimate of 50% is, however, likely to be a lower limit for the following reasons. First, the mechanisms underlying the GWAS loci are certainly not limited to the molecular phenotypes analyzed in the present study. Hence, adding other types of xQTLs (e.g., post-transcriptional modification and metabolite QTLs) is expected to link more GWAS loci with molecular intermediates and unveil new mechanisms. Second, the sample sizes of the xQTL studies are often small so that some of the small- to moderate-effect xQTLs remain undetected, as evidenced in the previous eQTL studies[22,55] where the number of genes detected with a *cis*-eQTL increased from 44% to 88% when the sample size increased from ~5,300 to ~32,000. In addition to the xQTL sample size, small sample sizes of GWAS can also lead to a decrease in OPERA power (Figure S26). Third, given that the xQTL data used in this study were predominantly from blood, xQTLs with tissue-, cell-type-, or developmental-stage-specific effects were likely to be under-represented, giving rise to another type of power loss. For example, the association between *MSMB* gene and prostate cancer was missed in the analysis with the blood eQTL data but identified with the prostate eQTL data. When the analyzed xQTL data are from bulk tissue, the power of xQTL detection can vary substantially across xQTL datasets from different studies because of differences in effect size, cellular specificity, cellular composition,

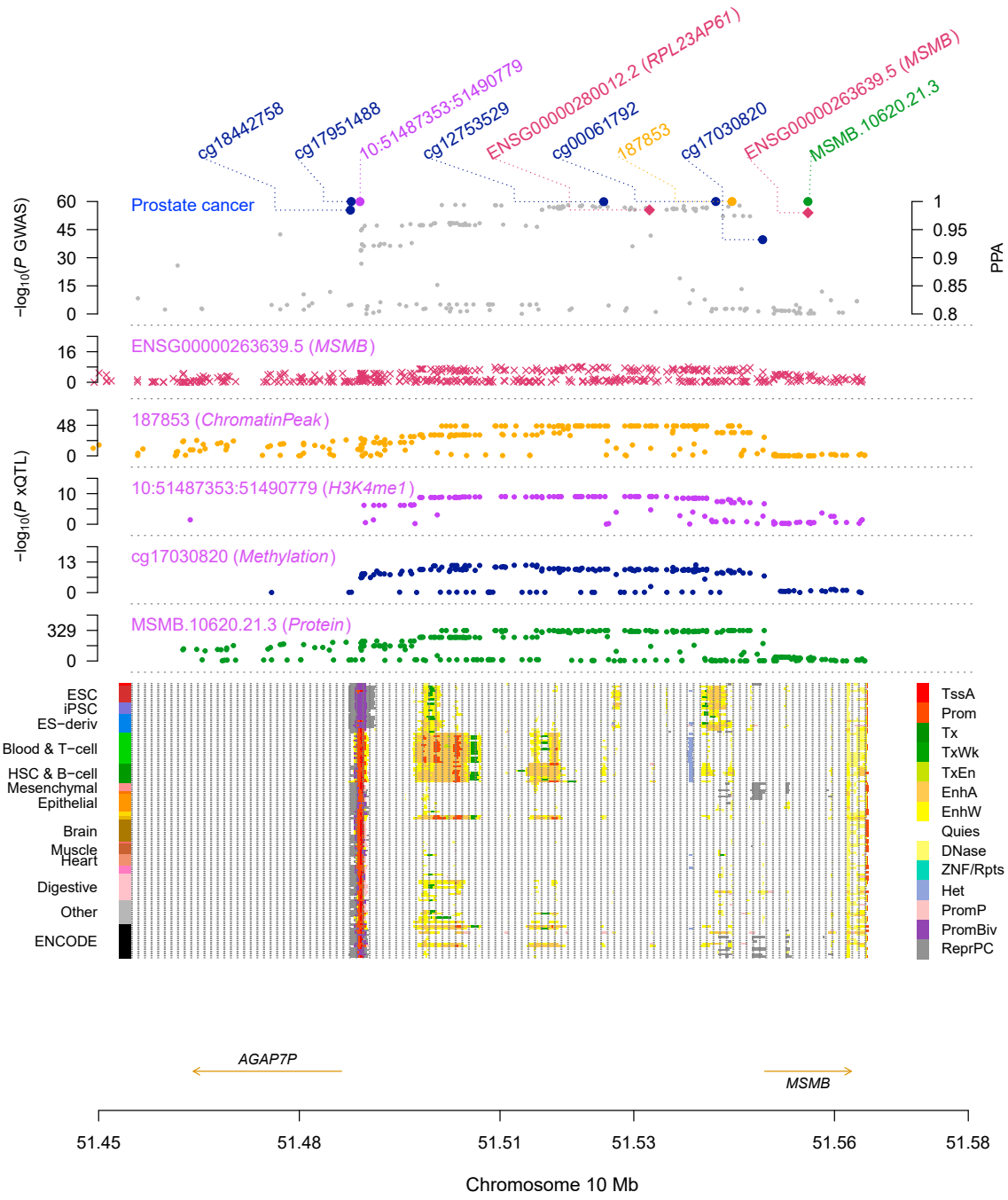


Figure 6. Prioritizing genes at the *MSMB* locus for prostate cancer

The top track shows the $-\log_{10}$(p values) of the GWAS SNPs (gray dots) for prostate cancer. The red diamonds represent OPERA marginal PPA for associations of genes using eQTL data from the GTEx prostate tissue, and circles with different colors are the marginal PPA from OPERA for associations of chromatin peak, histone modification, DNA methylation, and protein with the prostate cancer, respectively. The second track shows $-\log_{10}$(p values) of SNP-gene associations for *MSMB* gene from the GTEx prostate tissue. The subsequent tracks show $-\log_{10}$(p values) of SNP associations for other molecular phenotypes from the corresponding xQTL datasets (STAR Methods). The track on the bottom shows 14 chromatin state annotations inferred from the 127 Roadmap Epigenomics Mapping Consortium samples.

and/or sample size, which may explain some of the GWAS loci at which the GWAS signals are shared with only one type or a few types of xQTLs. Fourth, some of the omics layers had limited coverage; for example, for the proteome only 687 proteins were included in our analysis, largely owing to the limitation in the technology used to measure proteins. Data from more

advanced techniques with substantially higher coverage (e.g., protein abundance measures from mass spectrometry or DNA methylation measures from whole-genome bisulfite sequencing) are promising to provide a more comprehensive xQTL profile to saturate the GWAS loci. Fifth, we restricted our analyses to the molecular phenotypes within 1 Mb distance of the GWAS signal because of the complexity of the computation and data availability. Therefore, long-range genetic regulations beyond this window (including the *trans*-xQTLs) were not included in our analysis. Sixth, the small sample sizes for some molecular phenotypes could lead to conservative estimates of $\hat{\pi}$ for the alternative hypotheses (Figure S3), which might reduce the power to detect the true associations.

Limitations of the study

We note several limitations of our work. First, OPERA is a two-stage approach where we first estimated the proportions of the model configurations from independent xQTLs with balanced data and then used the estimates as priors to compute the posterior probabilities of associations at all xQTLs. While a linear relationship ($r = 0.66$) was observed between $\hat{\pi}_{PPA}$ from stage-1 analysis and proportion of GWAS loci explained by combinations in stage-2 analysis (Figure S27), a full Bayesian approach that simultaneously performs parameter estimation and a hypothesis test is more favorable but computationally infeasible under the current method framework. Second, even though our method is efficient, with the computation time independent of the number of SNPs and sample sizes, further optimizations are needed to analyze a large number of xQTL datasets (e.g., >10). For example, it is infeasible to apply the current OPERA analysis to analyze all 49 GTEx tissues because of the computational burden. Other methods, such as CAFEH,[56] are capable of jointly analyzing data from multiple tissues but are limited to a single omics layer. Third, the multi-exposure HEIDI test is conservative partly because of the removal of loci without enough SNPs in common between the GWAS and xQTL datasets and because of the potential heterogeneity in per-SNP sample size, especially when the summary statistics are from multi-cohort meta-analyses. The conservativeness of the multi-exposure HEIDI test also lies in the fact that it rejects associations for which the causal variants are not shared among exposures. For example, two exposures that are associated with the outcome through two distinct causal variants will not pass the multi-exposure HEIDI test if the causal variants are in LD, regardless of whether the effects of the exposures on the outcome are causative or not (Figure S28). Of note, if the causal variants for the exposures are independent, each exposure-outcome association can pass the single-exposure HEIDI test. Therefore, we reason that the exposure-outcome associations that pass the stage-2 analysis but fail in the multi-exposure HEIDI test could be worth considering in follow-up studies. Fourth, we interpret the detected associations from our method as pleiotropy; further analysis to distinguish causality (i.e., vertical pleiotropy) from horizontal pleiotropy requires multiple independent *trans*-xQTLs for a single molecular phenotype measure. Despite these limitations, our results revealed that more than half of GWAS signals shared with the molecular phenotypes and provide new insights into the biological mechanisms underpinning complex trait variation. To obtain a more saturated map linking molecular intermediates to GWAS loci requires xQTL data with more molecular phenotypes, measured while considering spatiotemporal effects in larger samples and potentially in different populations.

ETHICAL APPROVAL

This study was approved by the University of Queensland Human Research Ethics Committee B (approval no. 2011001173) and the Ethics Committee of Westlake University (approval no. 20200722YJ001).

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Ethical approval
- METHOD DETAILS
  - Single-exposure OPERA model
  - Multi-exposure OPERA model
  - Stage-1 OPERA analysis
  - Stage-2 OPERA analysis
  - Multi-exposure HEIDI test
  - Correcting for the cross-exposure leakage effect
  - Simulations based on imputed genotype data from the UK Biobank
  - Data used in this study

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.xgen.2023.100344.

ACKNOWLEDGMENTS

This research was supported by the "Pioneer" and "Leading Goose" R&D Program of Zhejiang (2022SDXHDX0001), the Leading Innovative and Entrepreneur Team Introduction Program of Zhejiang (2021R01013), the Australian National Health and Medical Research Council (1107258, 1113400, and 1177268), the Westlake Education Foundation, and the Research Center for Industries of the Future at Westlake University. We thank the Westlake University High-Performance Computing Center for assistance in computing. This study makes use of data from dbGaP (accessions: phs000428 and phs000424) and the UK Biobank (applications: 12505 and 66982).

AUTHOR CONTRIBUTIONS

J.Y. and J.Z. conceived and supervised the study. J.Z., Y.W., and J.Y. developed the methods and designed the experiment. Y.W. conducted all analyses and developed the software with the assistance or guidance from J.Y., J.Z., T.Q., N.R.W., and P.M.V. Y.W., J.Z., and J.Y. wrote the manuscript with the participation of all authors. All authors approved the final version of the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: October 3, 2022
Revised: April 4, 2023
Accepted: May 23, 2023
Published: June 19, 2023

REFERENCES

1. Visscher, P.M., Wray, N.R., Zhang, Q., Sklar, P., McCarthy, M.I., Brown, M.A., and Yang, J. (2017). 10 Years of GWAS discovery: biology, function, and translation. Am. J. Hum. Genet. *101*, 5–22. https://doi.org/10.1016/j.ajhg.2017.06.005.
2. Hormozdiari, F., Gazal, S., van de Geijn, B., Finucane, H.K., Ju, C.J.T., Loh, P.R., Schoech, A., Reshef, Y., Liu, X., O'Connor, L., et al. (2018). Leveraging molecular quantitative trait loci to understand the genetic architecture of diseases and complex traits. Nat. Genet. *50*, 1041–1047. https://doi.org/10.1038/s41588-018-0148-2.
3. Hannon, E., Weedon, M., Bray, N., O'Donovan, M., and Mill, J. (2017). Pleiotropic effects of trait-associated genetic variation on DNA methylation: utility for refining GWAS loci. Am. J. Hum. Genet. *100*, 954–959. https://doi.org/10.1016/j.ajhg.2017.04.013.
4. Gusev, A., Mancuso, N., Won, H., Kousi, M., Finucane, H.K., Reshef, Y., Song, L., Safi, A., Schizophrenia Working Group of the Psychiatric Genomics Consortium; and McCarroll, S., et al. (2018). Transcriptome-wide association study of schizophrenia and chromatin activity yields mechanistic disease insights. Nat. Genet. *50*, 538–548. https://doi.org/10.1038/s41588-018-0092-1.
5. Watanabe, K., Taskesen, E., van Bochoven, A., and Posthuma, D. (2017). Functional mapping and annotation of genetic associations with FUMA. Nat. Commun. *8*, 1826. https://doi.org/10.1038/s41467-017-01261-5.
6. Hormozdiari, F., van de Bunt, M., Segrè, A.V., Li, X., Joo, J.W.J., Bilow, M., Sul, J.H., Sankararaman, S., Pasaniuc, B., and Eskin, E. (2016). Colocalization of GWAS and eQTL signals detects target genes. Am. J. Hum. Genet. *99*, 1245–1260. https://doi.org/10.1016/j.ajhg.2016.10.003.
7. Richardson, T.G., Haycock, P.C., Zheng, J., Timpson, N.J., Gaunt, T.R., Davey Smith, G., Relton, C.L., and Hemani, G. (2018). Systematic Mendelian randomization framework elucidates hundreds of CpG sites which may mediate the influence of genetic variants on disease. Hum. Mol. Genet. *27*, 3293–3304. https://doi.org/10.1093/hmg/ddy210.
8. Zhu, Z., Zhang, F., Hu, H., Bakshi, A., Robinson, M.R., Powell, J.E., Montgomery, G.W., Goddard, M.E., Wray, N.R., Visscher, P.M., and Yang, J. (2016). Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat. Genet. *48*, 481–487. https://doi.org/10.1038/ng.3538.
9. Gusev, A., Ko, A., Shi, H., Bhatia, G., Chung, W., Penninx, B., Jansen, R., de Geus, E.J.C., Boomsma, D.I., Wright, F.A., et al. (2016). Integrative approaches for large-scale transcriptome-wide association studies. Nat. Genet. *48*, 245–252. https://doi.org/10.1038/ng.3506.
10. Giambartolomei, C., Vukcevic, D., Schadt, E.E., Franke, L., Hingorani, A.D., Wallace, C., and Plagnol, V. (2014). Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLoS Genet. *10*, e1004383. https://doi.org/10.1371/journal.pgen.1004383.
11. Gamazon, E.R., Wheeler, H.E., Shah, K.P., Mozaffari, S.V., Aquino-Michaels, K., Carroll, R.J., Eyler, A.E., Denny, J.C., GTEx Consortium; and Nicolae, D.L., et al. (2015). A gene-based association method for mapping traits using reference transcriptome data. Nat. Genet. *47*, 1091–1098. https://doi.org/10.1038/ng.3367.
12. Nica, A.C., Montgomery, S.B., Dimas, A.S., Stranger, B.E., Beazley, C., Barroso, I., and Dermitzakis, E.T. (2010). Candidate causal regulatory effects by integration of expression QTLs with complex trait genetic associations. PLoS Genet. *6*, e1000895. https://doi.org/10.1371/journal.pgen.1000895.
13. He, X., Fuller, C.K., Song, Y., Meng, Q., Zhang, B., Yang, X., and Li, H. (2013). Sherlock: detecting gene-disease associations by matching patterns of expression QTL and GWAS. Am. J. Hum. Genet. *92*, 667–680. https://doi.org/10.1016/j.ajhg.2013.03.022.
14. Wen, X., Pique-Regi, R., and Luca, F. (2017). Integrating molecular QTL data into genome-wide genetic association analysis: probabilistic assessment of enrichment and colocalization. PLoS Genet. *13*, e1006646. https://doi.org/10.1371/journal.pgen.1006646.
15. Gleason, K.J., Yang, F., Pierce, B.L., He, X., and Chen, L.S. (2020). Primo: integration of multiple GWAS and omics QTL summary statistics for elucidation of molecular mechanisms of trait-associated SNPs and detection of pleiotropy in complex traits. Genome Biol. *21*, 236. https://doi.org/10.1186/s13059-020-02125-w.
16. Wu, Y., Zeng, J., Zhang, F., Zhu, Z., Qi, T., Zheng, Z., Lloyd-Jones, L.R., Marioni, R.E., Martin, N.G., Montgomery, G.W., et al. (2018). Integrative analysis of omics summary data reveals putative mechanisms underlying complex traits. Nat. Commun. *9*, 918. https://doi.org/10.1038/s41467-018-03371-0.
17. Zheng, J., Haberland, V., Baird, D., Walker, V., Haycock, P.C., Hurle, M.R., Gutteridge, A., Erola, P., Liu, Y., Luo, S., et al. (2020). Phenome-wide Mendelian randomization mapping the influence of the plasma proteome on complex diseases. Nat. Genet. *52*, 1122–1131. https://doi.org/10.1038/s41588-020-0682-6.
18. Walker, R.L., Ramaswami, G., Hartl, C., Mancuso, N., Gandal, M.J., de la Torre-Ubieta, L., Pasaniuc, B., Stein, J.L., and Geschwind, D.H. (2019). Genetic control of expression and splicing in developing human brain informs disease mechanisms. Cell *179*, 750–771.e22. https://doi.org/10.1016/j.cell.2019.09.021.
19. Jian, Y., Ting, Q., Yang, W., Futao, Z., and Jian, Z. (2021). Research Square. https://doi.org/10.21203/rs.3.rs-155233/v1.
20. Giambartolomei, C., Zhenli Liu, J., Zhang, W., Hauberg, M., Shi, H., Boocock, J., Pickrell, J., Jaffe, A.E., CommonMind Consortium; Pasaniuc, B., and Roussos, P. (2018). A Bayesian framework for multiple trait colocalization from summary association statistics. Bioinformatics *34*, 2538–2545. https://doi.org/10.1093/bioinformatics/bty147.
21. Foley, C.N., Staley, J.R., Breen, P.G., Sun, B.B., Kirk, P.D.W., Burgess, S., and Howson, J.M.M. (2021). A fast and efficient colocalization algorithm for identifying shared genetic risk factors across multiple traits. Nat. Commun. *12*, 764. https://doi.org/10.1038/s41467-020-20885-8.
22. Võsa, U., Claringbould, A., Westra, H.-J., Bonder, M.J., Deelen, P., Zeng, B., Kirsten, H., Saha, A., Kreuzhuber, R., Yazar, S., et al. (2021). Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. Nat. Genet. *53*, 1300–1310. https://doi.org/10.1038/s41588-021-00913-z.
23. Sun, B.B., Maranville, J.C., Peters, J.E., Stacey, D., Staley, J.R., Blackshaw, J., Burgess, S., Jiang, T., Paige, E., Surendran, P., et al. (2018). Genomic atlas of the human plasma proteome. Nature *558*, 73–79. https://doi.org/10.1038/s41586-018-0175-2.
24. Kumasaka, N., Knights, A.J., and Gaffney, D.J. (2019). High-resolution genetic mapping of putative causal interactions between regions of open chromatin. Nat. Genet. *51*, 128–137. https://doi.org/10.1038/s41588-018-0278-6.
25. Chen, L., Ge, B., Casale, F.P., Vasquez, L., Kwan, T., Garrido-Martín, D., Watt, S., Yan, Y., Kundu, K., Ecker, S., et al. (2016). Genetic drivers of epigenetic and transcriptional variation in human immune cells. Cell *167*, 1398–1414.e24. https://doi.org/10.1016/j.cell.2016.10.026.
26. McRae, A.F., Marioni, R.E., Shah, S., Yang, J., Powell, J.E., Harris, S.E., Gibson, J., Henders, A.K., Bowdler, L., Painter, J.N., et al. (2018). Identification of 55,000 replicated DNA methylation QTL. Sci. Rep. *8*, 17605. https://doi.org/10.1038/s41598-018-35871-w.

27. Li, L., Huang, K.-L., Gao, Y., Cui, Y., Wang, G., Elrod, N.D., Li, Y., Chen, Y.E., Ji, P., Peng, F., et al. (2021). An atlas of alternative polyadenylation quantitative trait loci contributing to complex trait and disease heritability. Nat. Genet. *53*, 994–1005. https://doi.org/10.1038/s41588-021-00864-5.

28. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., and Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. *81*, 559–575. https://doi.org/10.1086/519795.

29. Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L.T., Sharp, K., Motyer, A., Vukcevic, D., Delaneau, O., O'Connell, J., et al. (2018). The UK Biobank resource with deep phenotyping and genomic data. Nature *562*, 203–209. https://doi.org/10.1038/s41586-018-0579-z.

30. Stephens, M. (2017). False discovery rates: a new deal. Biostatistics *18*, 275–294. https://doi.org/10.1093/biostatistics/kxw041.

31. Storey, J.D., and Tibshirani, R. (2003). Statistical significance for genome-wide studies. Proc. Natl. Acad. Sci. USA *100*, 9440–9445. https://doi.org/10.1073/pnas.1530509100.

32. Fernando, R., Toosi, A., Wolc, A., Garrick, D., and Dekkers, J. (2017). Application of whole-genome prediction methods for genome-wide association studies: a bayesian approach. J. Agric. Biol. Environ. Stat. *22*, 172–193. https://doi.org/10.1007/s13253-017-0277-6.

33. Yengo, L., Sidorenko, J., Kemper, K.E., Zheng, Z., Wood, A.R., Weedon, M.N., Frayling, T.M., Hirschhorn, J., Yang, J., and Visscher, P.M.; GIANT Consortium (2018). Meta-analysis of genome-wide association studies for height and body mass index in ~700000 individuals of European ancestry. Hum. Mol. Genet. *27*, 3641–3649. https://doi.org/10.1093/hmg/ddy271.

34. Astle, W.J., Elding, H., Jiang, T., Allen, D., Ruklisa, D., Mann, A.L., Mead, D., Bouman, H., Riveros-Mckay, F., Kostadima, M.A., et al. (2016). The allelic landscape of human blood cell trait variation and links to common complex disease. Cell *167*, 1415–1429.e19. https://doi.org/10.1016/j.cell.2016.10.042.

35. Xue, A., Wu, Y., Zhu, Z., Zhang, F., Kemper, K.E., Zheng, Z., Yengo, L., Lloyd-Jones, L.R., Sidorenko, J., Wu, Y., et al. (2018). Genome-wide association analyses identify 143 risk variants and putative regulatory mechanisms for type 2 diabetes. Nat. Commun. *9*, 2941. https://doi.org/10.1038/s41467-018-04951-w.

36. Nikpay, M., Goel, A., Won, H.-H., Hall, L.M., Willenborg, C., Kanoni, S., Saleheen, D., Kyriakou, T., Nelson, C.P., Hopewell, J.C., et al. (2015). A comprehensive 1000 Genomes–based genome-wide association meta-analysis of coronary artery disease. Nat. Genet. *47*, 1121–1130. https://doi.org/10.1038/ng.3396.

37. Willer, C.J., Schmidt, E.M., Sengupta, S., Peloso, G.M., Gustafsson, S., Kanoni, S., Ganna, A., Chen, J., Buchkovich, M.L., Mora, S., et al. (2013). Discovery and refinement of loci associated with lipid levels. Nat. Genet. *45*, 1274–1283. https://doi.org/10.1038/ng.2797.

38. Liu, J.Z., van Sommeren, S., Huang, H., Ng, S.C., Alberts, R., Takahashi, A., Ripke, S., Lee, J.C., Jostins, L., Shah, T., et al. (2015). Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. Nat. Genet. *47*, 979–986. https://doi.org/10.1038/ng.3359.

39. Jansen, I.E., Savage, J.E., Watanabe, K., Bryois, J., Williams, D.M., Steinberg, S., Sealock, J., Karlsson, I.K., Hägg, S., Athanasiu, L., et al. (2019). Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. Nat. Genet. *51*, 404–413. https://doi.org/10.1038/s41588-018-0311-9.

40. Pardiñas, A.F., Holmans, P., Pocklington, A.J., Escott-Price, V., Ripke, S., Carrera, N., Legge, S.E., Bishop, S., Cameron, D., Hamshere, M.L., et al. (2018). Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. Nat. Genet. *50*, 381–389. https://doi.org/10.1038/s41588-018-0059-2.

41. Lee, J.J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., Nguyen-Viet, T.A., Bowers, P., Sidorenko, J., Karlsson Linnér, R., et al. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. Nat. Genet. *50*, 1112–1121. https://doi.org/10.1038/s41588-018-0147-3.

42. Schumacher, F.R., Al Olama, A.A., Berndt, S.I., Benlloch, S., Ahmed, M., Saunders, E.J., Dadaev, T., Leongamornlert, D., Anokian, E., Cieza-Borrella, C., et al. (2018). Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. Nat. Genet. *50*, 928–936. https://doi.org/10.1038/s41588-018-0142-8.

43. Michailidou, K., Lindström, S., Dennis, J., Beesley, J., Hui, S., Kar, S., Lemaçon, A., Soucy, P., Glubb, D., Rostamianfar, A., et al. (2017). Association analysis identifies 65 new breast cancer risk loci. Nature *551*, 92–94. https://doi.org/10.1038/nature24284.

44. Lloyd-Jones, L.R., Holloway, A., McRae, A., Yang, J., Small, K., Zhao, J., Zeng, B., Bakshi, A., Metspalu, A., Dermitzakis, M., et al. (2017). The genetic architecture of gene expression in peripheral blood. Am. J. Hum. Genet. *100*, 371. https://doi.org/10.1016/j.ajhg.2017.01.026.

45. GTEx Consortium (2020). The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science *369*, 1318–1330. https://doi.org/10.1126/science.aaz1776.

46. Yang, J., Ferreira, T., Morris, A.P., Medland, S.E., Genetic Investigation of ANthropometric Traits GIANT Consortium; DIAbetes Genetics Replication And Meta-analysis DIAGRAM Consortium; Madden, P.A.F., Heath, A.C., Martin, N.G., Montgomery, G.W., et al. (2012). Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nat. Genet. *44*, 369–375. https://doi.org/10.1038/ng.2213.

47. Wood, A.R., Esko, T., Yang, J., Vedantam, S., Pers, T.H., Gustafsson, S., Chu, A.Y., Estrada, K., Luan, J., Kutalik, Z., et al. (2014). Defining the role of common variation in the genomic and biological architecture of adult human height. Nat. Genet. *46*, 1173–1186. https://doi.org/10.1038/ng.3097.

48. Locke, A.E., Kahali, B., Berndt, S.I., Justice, A.E., Pers, T.H., Day, F.R., Powell, C., Vedantam, S., Buchkovich, M.L., Yang, J., et al. (2015). Genetic studies of body mass index yield new insights for obesity biology. Nature *518*, 197–206. https://doi.org/10.1038/nature14177.

49. Folkersen, L., Gustafsson, S., Wang, Q., Hansen, D.H., Hedman, Å.K., Schork, A., Page, K., Zhernakova, D.V., Wu, Y., Peters, J., et al. (2020). Genomic and drug target evaluation of 90 cardiovascular proteins in 30,931 individuals. Nat. Metab. *2*, 1135–1148. https://doi.org/10.1038/s42255-020-00287-2.

50. Bradshaw, E.M., Chibnik, L.B., Keenan, B.T., Ottoboni, L., Raj, T., Tang, A., Rosenkrantz, L.L., Imboywa, S., Lee, M., Von Korff, A., et al. (2013). CD33 Alzheimer's disease locus: altered monocyte function and amyloid biology. Nat. Neurosci. *16*, 848–850. https://doi.org/10.1038/nn.3435.

51. Interleukin-6 Receptor Mendelian Randomisation Analysis IL6R MR Consortium; Swerdlow, D.I., Holmes, M.V., Kuchenbaecker, K.B., Engmann, J.E.L., Shah, T., Sofat, R., Guo, Y., Chung, C., Peasey, A., Pfister, R., et al. (2012). The Interleukin-6 Receptor Mendelian Randomisation Analysis (IL6R MR) Consortium (2012). The interleukin-6 receptor as a target for prevention of coronary heart disease: a mendelian randomisation analysis. Lancet *379*, 1214–1224. https://doi.org/10.1016/S0140-6736(12)60110-X.

52. Chun, S., Casparino, A., Patsopoulos, N.A., Croteau-Chonka, D.C., Raby, B.A., De Jager, P.L., Sunyaev, S.R., and Cotsapas, C. (2017). Limited statistical evidence for shared genetic effects of eQTLs and autoimmune-disease-associated loci in three major immune-cell types. Nat. Genet. *49*, 600–605. https://doi.org/10.1038/ng.3795.

53. GTEx Consortium (2015). The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science *348*, 648–660. https://doi.org/10.1126/science.1262110.

54. Ming, J., Wang, T., and Yang, C. (2020). LPM: a latent probit model to characterize the relationship among complex traits using summary statistics from multiple GWASs and functional annotations. Bioinformatics *36*, 2506–2514. https://doi.org/10.1093/bioinformatics/btz947.

55. Westra, H.-J., Peters, M.J., Esko, T., Yaghootkar, H., Schurmann, C., Kettunen, J., Christiansen, M.W., Fairfax, B.P., Schramm, K., Powell, J.E., et al. (2013). Systematic identification of trans eQTLs as putative drivers of known disease associations. Nat. Genet. *45*, 1238–1243. https://doi.org/10.1038/ng.2756.

56. Arvanitis, M., Tayeb, K., Strober, B.J., and Battle, A. (2022). Redefining tissue specificity of genetic regulation of gene expression in the presence of allelic heterogeneity. Am. J. Hum. Genet. *109*, 223–239. https://doi.org/10.1016/j.ajhg.2022.01.002.

57. Qi, T., Wu, Y., Zeng, J., Zhang, F., Xue, A., Jiang, L., Zhu, Z., Kemper, K., Yengo, L., Zheng, Z., eQTLGen Consortium; Marioni, R.E., Montgomery, G.W., Deary, I.J., Wray, N.R., Visscher, P.M., McRae, A.F., and Yang, J. (2018). Identifying gene targets for brain-related traits using transcriptomic and methylomic data from blood. Nat. Commun. *9*, 2282. https://doi.org/10.1038/s41467-018-04558-1.

58. Kuonen, D. (1999). Saddlepoint approximations for distributions of quadratic forms in normal variables. Biometrika *86*, 929–935.

59. Lawlor, D.A., Harbord, R.M., Sterne, J.A.C., Timpson, N., and Davey Smith, G. (2008). Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Stat. Med. *27*, 1133–1163. https://doi.org/10.1002/sim.3034.

60. Sonnega, A., Faul, J.D., Ofstedal, M.B., Langa, K.M., Phillips, J.W.R., and Weir, D.R. (2014). Cohort profile: the health and retirement study (HRS). Int. J. Epidemiol. *43*, 576–585. https://doi.org/10.1093/ije/dyu067.

61. Chen, W., Wu, Y., Zheng, Z., Qi, T., Visscher, P.M., Zhu, Z., and Yang, J. (2021). Improved analyses of GWAS summary statistics by reducing data heterogeneity and errors. Nat. Commun. *12*, 7117. https://doi.org/10.1038/s41467-021-27438-7.

62. Yang, J., Lee, S.H., Goddard, M.E., and Visscher, P.M. (2011). GCTA: a tool for genome-wide complex trait analysis. Am. J. Hum. Genet. *88*, 76–82. https://doi.org/10.1016/j.ajhg.2010.11.011.

63. Berisa, T., and Pickrell, J.K. (2016). Approximately independent linkage disequilibrium blocks in human populations. Bioinformatics *32*, 283–285. https://doi.org/10.1093/bioinformatics/btv546.

64. Zeng, J., Xue, A., Jiang, L., Lloyd-Jones, L.R., Wu, Y., Wang, H., Zheng, Z., Yengo, L., Kemper, K.E., Goddard, M.E., et al. (2021). Widespread signatures of natural selection across human complex traits and functional genomic categories. Nat. Commun. *12*, 1164. https://doi.org/10.1038/s41467-021-21446-3.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
UK Biobank data	Bycroft et al.[29]	http://www.ukbiobank.ac.uk
Cis-eQTL summary statistics	Võsa et al.[22]	https://www.eqtlgen.org/cis-eqtls.html
Cis-hQTL summary statistics	Chen et al.[25]	ftp://ftp.ebi.ac.uk/pub/databases/blueprint/blueprint_Epivar/qtl_as/
Cis-mQTL summary statistics	Wu et al.[16]	https://yanglab.westlake.edu.cn/software/smr/#DataResource Zenodo https://doi.org/10.5281/zenodo.7949311
Software and algorithms		
OPERA	This study	https://github.com/wuyangf7/OPERA Zenodo https://doi.org/10.5281/zenodo.7903457
GCTA	Yang et al.[57]	https://yanglab.westlake.edu.cn/software/gcta/#Download
SMR	Zhu et al.[8]	https://yanglab.westlake.edu.cn/software/smr/
MOLOC	Giambartolomei et al.[20]	https://github.com/clagiamba/moloc
Primo	Gleason et al.[15]	https://github.com/primocms/primo
HyPrColoc	Foley et al.[21]	https://github.com/jrs95/hyprcoloc

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Jian Yang (jian.yang@westlake.edu.cn).

Materials availability

This study did not generate new unique reagents.

Data and code availability

The source code of OPERA is available at https://github.com/wuyangf7/OPERA or Zenodo: https://zenodo.org/record/7903457#.ZFbmVOzMKQc (https://doi.org/10.5281/zenodo.7903457). The summary-level *cis*-eQTL data of the eQTLGen project are publicly available at https://www.eqtlgen.org/cis-eqtls.html. The blueprint data are available at ftp://ftp.ebi.ac.uk/pub/databases/blueprint/blueprint_Epivar/qtl_as/. The summary statistics of the mQTL data are available at https://zenodo.org/record/7949311#.ZGa0POzMKog (https://doi.org/10.5281/zenodo.7951839) or https://yanglab.westlake.edu.cn/software/smr/#DataResource. The HRS data are available from dbGap (accession number: phs000428) and the UK Biobank data are available through formal application to the UK Biobank (http://www.ukbiobank.ac.uk). All the other datasets used in this study are available in the public domain.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Ethical approval

This study was approved by the University of Queensland Human Research Ethics Committee B (approval no. 2011001173) and the Ethics Committee of Westlake University (approval no. 20200722YJ001).

METHOD DETAILS

Single-exposure OPERA model

OPERA is a Bayesian method to test for associations of multiple molecular phenotypes (exposures) simultaneously with a complex trait (outcome). For the ease of illustration, we will start describing the OPERA model with a single exposure and then generalise it to a multi-exposure model.

When only one exposure is considered, OPERA becomes a Bayesian analysis of the SMR model.[8] SMR is a method that uses a genetic variant as an instrument variable to test for association between a molecular phenotype (exposure) and a complex trait (outcome). If we denote z as an instrument SNP associated with the exposure x, and y as the outcome, the SMR estimate of the effect

of x on y is $\widehat{b}_{xy} = \widehat{b}_{zy}/\widehat{b}_{zx}$ with $\mathrm{var}(\widehat{b}_{xy}) \approx \frac{b_{zy}^2}{b_{zx}^2}\left(\frac{\mathrm{var}(\widehat{b}_{zx})}{b_{zx}^2} + \frac{\mathrm{var}(\widehat{b}_{zy})}{b_{zy}^2}\right)$, assuming that $\widehat{b}_{zx}$ (the xQTL effect) and $\widehat{b}_{zy}$ (the GWAS effect) are estimated from independent samples. We model $\widehat{b}_{xy}$ as (for simplicity, b_{xy} will be denoted by b hereafter)

$$\widehat{b} = b + e$$

where b is the true effect of the exposure on the outcome and e is the estimation error with $e \sim N(0, \sigma_e^2)$. We assume that b follows a point-normal mixture distribution

$$b \sim \pi N(0, \sigma_b^2) + (1 - \pi)\varphi$$

with $\boldsymbol{\pi}$ being the proportion of sites (e.g., genes or DNA methylation sites) of the exposure with nonzero effects and φ being the point mass at zero. In this case, $\widehat{b}$ follows a mixture normal distribution

$$\widehat{b} \sim \pi N(0, \sigma_e^2 + \sigma_b^2) + (1 - \pi)N(0, \sigma_e^2)$$

We estimate σ_e^2 from the standard error of $\widehat{b}$ reported by SMR and estimate σ_b^2 (i.e., variance of the non-zero effects) by the variance of the estimated SMR effects at FDR <0.05, adjusting for the estimation errors.[57] Although we specifically used $\widehat{b}$ from SMR that uses the top xQTL as an instrument, the OPERA framework can be extended to use $\widehat{b}$ estimated based on multiple instrumental SNPs.

Given the models above, the marginal likelihood of the alternative ($H_1 : b \neq 0$) and null ($H_0 : b = 0$) hypotheses are

$$f(\widehat{b}|H_1) \propto \left(C\sigma_b^2\right)^{-\frac{1}{2}} exp\left\{\frac{\left(C^{-1} - 1\right)\widehat{b}^2}{2\sigma_e^2}\right\} \quad \text{(Equation 1)}$$

$$f(\widehat{b}|H_0) \propto \left(\sigma_e^2\right)^{-\frac{1}{2}} exp\left\{-\frac{\widehat{b}^2}{2\sigma_e^2}\right\} \quad \text{(Equation 2)}$$

where $C = 1 + \frac{\sigma_e^2}{\sigma_b^2}$. Then, the posterior probability supporting H_1 given $\widehat{b}$ is

$$\Pr(H_1|\widehat{b}) = \frac{f(\widehat{b}|H_1) \times \pi}{f(\widehat{b}|H_0) \times (1 - \pi) + f(\widehat{b}|H_1) \times \pi}$$

Multi-exposure OPERA model

The single-exposure OPERA model above can be extended to analyze multiple types of xQTL summary data jointly with the GWAS summary data by regarding the different types of molecular phenotypes as multiple exposures (Figure 1). If there are t molecular phenotypes in a genomic region with one site for each molecular phenotype (i.e., t exposures), then $\widehat{\boldsymbol{b}} = [\widehat{b}_1, \ldots, \widehat{b}_t]$ is a vector of the SMR estimated effects of the exposures on the outcome. For exposure i, we introduce a dummy variable $\lambda_i \in \{0,1\}$ to indicate whether this exposure affects the outcome (i.e., if $\lambda_i = 0, b_i = 0$; if $\lambda_i = 1, b_i \neq 0$). Hence, there are $\Omega = 2^t$ possible patterns of how the t exposures are associated with the outcome (referred to as "configurations"; see Figure 1 for example). Let δ be a random variable that takes values of $1, 2, \ldots, \Omega$ to index the configurations with probabilities $\boldsymbol{\pi} = [\pi_1, \pi_2, \ldots, \pi_\Omega]$, i.e.,

$$\delta = \begin{cases} 1, & \pi_1 \\ \vdots & \vdots \\ \Omega, & \pi_\Omega \end{cases}$$

The posterior probability for supporting the γ^{th} configuration (PPC_γ) is

$$PPC_\gamma = \Pr(\delta = \gamma|\widehat{\boldsymbol{b}}, \boldsymbol{\pi}) = \frac{f(\widehat{\boldsymbol{b}}|\delta = \gamma) \times f(\delta = \gamma|\pi_\gamma)}{f(\widehat{\boldsymbol{b}}|\boldsymbol{\pi})} = \frac{f(\widehat{\boldsymbol{b}}|\delta = \gamma) \times \pi_\gamma}{\sum_{k=1}^{\Omega} f(\widehat{\boldsymbol{b}}|\delta = k) \times \pi_k} \quad \text{(Equation 3)}$$

where $f(\widehat{\boldsymbol{b}}|\delta = \gamma)$ is the joint likelihood of data across t exposures under configuration γ.

To compute PPC_γ, we need to know $f(\widehat{\boldsymbol{b}}|\delta = \gamma)$ and $\boldsymbol{\pi}$ (i.e., the global proportions of association patterns of the t molecular phenotypes with the complex trait). Assuming that the underlying true effects (i.e., $\boldsymbol{b}$) are independent among molecular phenotypes and the xQTL data are obtained from independent samples, $f(\widehat{\boldsymbol{b}}|\delta = \gamma)$ can be computed as the product of the marginal likelihood for each exposure,

$$f(\widehat{\boldsymbol{b}}|\delta = \gamma) = \prod_{i=1}^{t} f(\widehat{b}_i|H_1)^{\lambda_{i|\delta=\gamma}} f(\widehat{b}_i|H_0)^{1-\lambda_{i|\delta=\gamma}} \quad \text{(Equation 4)}$$

where λ_i is known given $\delta = \gamma, f(\hat{b}_i|H_1)$ and $f(\hat{b}_i|H_0)$ are the marginal likelihood for $\hat{b}_i$ under H_1 (i.e., $b_i \neq 0$) and under H_0 (i.e., $b_i = 0$), respectively (see Equations 1 and 2 above). When the xQTL data are from overlapping samples, we compute the likelihood from the multivariate normal distribution,

$$f(\hat{\boldsymbol{b}}|\delta = \gamma) = (2\pi|\boldsymbol{\Sigma}_b+\boldsymbol{\Sigma}_e|)^{-\frac{1}{2}} exp\left\{-\frac{1}{2}\hat{\boldsymbol{b}}^T[\boldsymbol{\Sigma}_b+\boldsymbol{\Sigma}_e]^{-1}\hat{\boldsymbol{b}}\right\}$$

where $\boldsymbol{\Sigma}_b$ is a diagonal matrix with the diagonal elements of $\sigma^2_{b_i}$ when $\lambda_i = 1$ and 0 otherwise, and $\boldsymbol{\Sigma}_e$ is the variance-covariance matrix of the estimation errors, ij^{th} element of which is $\rho_{ij}\sqrt{\sigma^2_{\hat{e}_i}\sigma^2_{\hat{e}_j}}$, with ρ_{ij} being the between study correlation due to sample overlap. We estimate ρ_{ij} using the SMR estimated effects of null exposure sites ($p > 0.01$) by a summary data-based approach described in Qi et al.[57] Computing the PPC_γ requires $\boldsymbol{\pi}$ (Equation 3), which are unknown in practice. Ideally, the problem can be solved by a full Bayesian analysis that makes simultaneous inference on the global parameters $\boldsymbol{\pi}$ and the local variables $\boldsymbol{\delta}$. However, some molecular phenotypes often have multiple sites or are missing at some loci, which makes it difficult to perform the full Bayesian analysis due to a heavy computation burden and complex correlation structures of molecular phenotype sites. We therefore propose a two-stage model to first estimate $\boldsymbol{\pi}$ using genome-wide quasi-independent loci with an uninformative prior distribution, and then compute PPC for each set of molecular phenotype sites at each genomic locus of interest, given the estimated $\boldsymbol{\pi}$ from the stage-1 analysis.

Stage-1 OPERA analysis

The aim of this analysis is to estimate $\boldsymbol{\pi}$ from data across the genome. Based on the molecular phenotype with the lowest genomic coverage, we select m approximately independent genomic loci (distance >500Kb between loci), each with t molecular phenotypes (note: we randomly sample one site as a representative if an exposure has multiple sites at a locus). We assume that $\boldsymbol{\pi} = \{\pi_1, \ldots, \pi_\Omega\}$ follows a Dirichlet distribution with parameters $\boldsymbol{\alpha} = \{\alpha_1, \ldots, \alpha_\Omega\}$,

$$f(\boldsymbol{\pi}) \propto \prod_{\gamma=1}^{\Omega} \pi_\gamma^{\alpha_\gamma - 1}$$

At locus l, δ_l has a categorical distribution given $\boldsymbol{\pi}$,

$$f(\delta_l|\boldsymbol{\pi}) = \prod_{\gamma=1}^{\Omega} \pi_\gamma^{|\delta_l = \gamma|}$$

where $|\delta_l = \gamma|$ equals to 1 when $\delta_l = \gamma$ and 0 otherwise. Therefore, the joint distribution of the data $\hat{\mathbf{B}} = \{\hat{\boldsymbol{b}}_1, \ldots, \hat{\boldsymbol{b}}_m\}$ and the unknowns $\boldsymbol{\delta}$ and $\boldsymbol{\pi}$, across m independent loci, is

$$f(\hat{\mathbf{B}}, \boldsymbol{\delta}, \boldsymbol{\pi}) = f(\hat{\mathbf{B}}|\boldsymbol{\delta}, \boldsymbol{\pi}) f(\boldsymbol{\delta}|\boldsymbol{\pi}) f(\boldsymbol{\pi}) = \prod_{l=1}^{m} f(\hat{\boldsymbol{b}}_l|\delta_l) f(\delta_l|\boldsymbol{\pi}) f(\boldsymbol{\pi})$$

where $f(\hat{\boldsymbol{b}}_l|\delta_l)$ is defined as in Equation 4 and $f(\hat{\mathbf{B}}|\boldsymbol{\delta}, \boldsymbol{\pi}) = \prod_{l=1}^{m} f(\hat{\boldsymbol{b}}_l|\delta_l)$, because loci are assumed to be independent and $\hat{\boldsymbol{b}}_l$ do not depend on $\boldsymbol{\pi}$ once conditional on δ_l. To make posterior inference on $\boldsymbol{\pi}$, the Gibbs sampling algorithm is employed to draw samples from the full conditional distributions for $\boldsymbol{\delta}$ and $\boldsymbol{\pi}$, iteratively as below.

Since all loci included in this stage are assumed to be independent, the full conditional distribution for δ_l only depends on $\hat{\boldsymbol{b}}_l$ but not $\hat{\boldsymbol{b}}$ at other loci.

$f(\delta_l|\hat{\boldsymbol{b}}_l, \boldsymbol{\pi}) = \frac{f(\hat{\boldsymbol{b}}_l|\delta_l) f(\delta_l|\boldsymbol{\pi})}{f(\hat{\boldsymbol{b}}_l, \boldsymbol{\pi})}$Conditional on $\hat{\boldsymbol{b}}_l$ and $\boldsymbol{\pi}$, the probability of $\delta_l = \gamma$ has the same form as Equation 3, i.e.,

$$\Pr(\delta_l = \gamma|\hat{\boldsymbol{b}}_l, \boldsymbol{\pi}) = \frac{f(\hat{\boldsymbol{b}}_l|\delta_l = \gamma) \times \pi_\gamma}{\sum_{k=1}^{\Omega} f(\hat{\boldsymbol{b}}_l|\delta_l = k) \times \pi_k}$$

based on which δ_l is sampled from a categorical distribution. The full conditional distribution for $\boldsymbol{\pi}$, given $\hat{\mathbf{B}}$ and $\boldsymbol{\delta}$, is

$$f(\boldsymbol{\pi}|\hat{\mathbf{B}}, \boldsymbol{\delta}) = \frac{f(\hat{\mathbf{B}}|\boldsymbol{\delta}) f(\boldsymbol{\delta}|\boldsymbol{\pi}) f(\boldsymbol{\pi})}{f(\hat{\mathbf{B}}|\boldsymbol{\delta}) f(\boldsymbol{\delta})} \propto f(\boldsymbol{\delta}|\boldsymbol{\pi}) f(\boldsymbol{\pi}) = \prod_{l=1}^{m} \prod_{\gamma=1}^{\Omega} \pi_\gamma^{|\delta_l = \gamma| + \alpha_\gamma - 1}$$

As mentioned above, $f(\hat{\mathbf{B}}|\boldsymbol{\delta})$ does not involve $\boldsymbol{\pi}$ and is therefore canceled out in the equation (namely, $\boldsymbol{\pi}$ connect to $\hat{\mathbf{B}}$ only through $\boldsymbol{\delta}$). The full conditional distribution for $\boldsymbol{\pi}$ is also a Dirichlet distribution with parameters being $(\boldsymbol{q} + \boldsymbol{\alpha})$, where $\boldsymbol{q} \in \{q_1, \ldots, q_\Omega\}$ with $q_\gamma = \sum_{l=1}^{m} |\delta_l = \gamma|$ being the number of loci consistent with configuration γ, and $\sum_{\gamma=1}^{\Omega} q_\gamma = m$. In this study, we assign a small, equal value to all α (i.e., $\alpha_1 = \ldots = \alpha_\Omega = 0.1$) to minimise the influence of prior specification on the posterior distribution of $\boldsymbol{\pi}$. We initialize $\boldsymbol{\pi} = \mathbf{1}/\Omega$ and generate 10,000 posterior samples. We discard the first 2,000 samples as burn-in and use the posterior mean over the preserved 8,000 posterior samples to estimate $\boldsymbol{\pi}$.

Stage-2 OPERA analysis

In the stage-2 analysis, we focus on the GWAS loci, with the aim to identify molecular phenotypes and their combinations that are associated with the complex trait. For ease of computation, we first remove exposure sites with $P_{SMR} > 0.05$, which are unlikely to be associated with the outcome. We enumerate all possible combinations of molecular phenotype sites (with one site for each molecular phenotype at a time) at a GWAS locus. The number of distinct OPERA analyses at a locus is equal to the product of the number of sites across all the molecular phenotype at the locus. For each combination of exposures at a locus, we compute the posterior probability supporting configuration γ as in Equation 3:

$$PPC_{\gamma} = \frac{f(\widehat{\boldsymbol{b}}|\delta = \gamma)\widehat{\pi}_{\gamma}}{\sum_{k=1}^{\Omega} f(\widehat{\boldsymbol{b}}|\delta = k)\widehat{\pi}_{k}}$$

with $\widehat{\boldsymbol{\pi}}$ obtained from the stage-1 analysis. We found that PPC could lead to false classifications when the effects of some exposures are too small to be detected. Take a combination with 2 causal exposures as an example (Figure S29). If $\delta_4 = [1,1]'$ is the causal configuration but the effect size of the second exposure is very small, the configuration with the highest PPC is likely to be $\delta_2 = [1,0]'$. In this case, the PPC corresponding to $\delta_2 = [1,0]'$ is therefore inflated because its true value is 0. However, neither the marginal PPA for one exposure nor the joint PPA for two exposures would be inflated. In this example, the marginal PPA for the second exposure (i.e., $PPA_{\lambda_2 = 1}$) and joint PPA for both exposure associations (i.e., $PPA_{\lambda_{1,2} = 1}$) are low, but the marginal PPA for the first exposure (i.e., $PPA_{\lambda_1 = 1}$) is still high (Figure S29). Thus, we report PPA as the final result and have evaluated the property of PPA by simulations (Figures 2B–2G). The null hypothesis of the OPERA test is that $\boldsymbol{b} = \boldsymbol{0}$ and the alternative hypothesis of the OPERA test is that any $b_i \neq 0$ (marginal PPA) or a subset of $\boldsymbol{b}$ are non-zero, e.g., $\boldsymbol{b_c} \neq 0$ (joint PPA). The marginal PPA for a single exposure i is computed as the sum of PPCs for configurations involving $\lambda_i = 1$ (denoted as $\Psi(\lambda_i = 1)$),

$$\text{PPA}_{m}\text{arginal} = \Pr(\lambda_i = 1|\widehat{\boldsymbol{b}}, \widehat{\pi}) = \sum_{\gamma \in \Psi(\lambda_i = 1)} \text{PPC}_{\gamma}$$

The PPA for the joint association of multiple exposures is computed as the sum of PPC for configurations involving $\boldsymbol{\lambda_c} = \boldsymbol{1}$ (denoted as $\Psi(\boldsymbol{\lambda_c} = \boldsymbol{1})$), where $\boldsymbol{c}$ is a vector of the exposures under consideration.

$$\text{PPA}_{\text{joint}} = \Pr(\boldsymbol{\lambda_c} = \boldsymbol{1}|\widehat{\boldsymbol{b}}, \widehat{\pi}) = \sum_{\gamma \in \Psi(\boldsymbol{\lambda_c} = \boldsymbol{1})} \text{PPC}_{\gamma}$$

Note that $\text{PPA}_{\text{joint}}$ is a joint probability of associations for the exposures under consideration but marginal to the other exposures. For instance, as illustrated in Figure S29, $\text{PPA}_1 = \text{PPC}_3 + \text{PPC}_4$, $\text{PPA}_2 = \text{PPC}_2 + \text{PPC}_4$ and $\text{PPA}_{1,2} = \text{PPC}_4$. When there are three exposures, $\text{PPA}_{1,2}$ is a sum of PPCs over configurations where both exposures 1 and 2 are associated regardless of whether exposure 3 is associated.

We claim significant associations for any single site or any combination of sites with PPA >0.9, which has been shown to be sufficient to control FDR below 0.05 in simulations (Figure 2C). To filter out associations not due to pleiotropy, we further apply a multi-exposure HEIDI test (see below). Using an empirical Bayesian approach,[30,32] we estimate the FDR given a PPA threshold of θ and HEIDI threshold of 0.01,

$$\widehat{\text{FDR}}(\theta) = \text{E}[(1 - \text{PPA}_i)|\text{PPA}_i > \theta \,\&\, P_{HEIDI,i} > 0.01] = \frac{\sum_i[(1 - \text{PPA}_i)|(\text{PPA}_i > \theta \,\&\, P_{HEIDI,i} > 0.01)]}{\#\{\text{PPA}_i > \theta \,\&\, P_{HEIDI,i} > 0.01\}} \quad \text{(Equation 5)}$$

where θ is the PPA threshold, and PPA_i and $P_{HEIDI,i}$ are the PPA and HEIDI test p-value for exposure i, respectively. The false positive rate given a PPA threshold of θ can be estimated as

$$\widehat{\text{FPR}}(\theta) = \frac{\sum_i[(1 - \text{PPA}_i)|(\text{PPA}_i > \theta \,\&\, P_{HEIDI,i} > 0.01)]}{M\widehat{\pi}_{null}} \quad \text{(Equation 6)}$$

where M is the total number of tests, and $\widehat{\pi}_{null}$ is the estimated proportion of null from the stage-1 analysis.

Multi-exposure HEIDI test

The associations identified from the Bayesian analysis above can occur even if the causal variants for the molecular phenotypes are distinct from, but in LD with, the causal variants for the trait. We have previously referred to this scenario as a linkage model and developed a method, HEIDI, to filter out exposure-outcome associations due to the linkage model.[8] In this study, we generalize the HEIDI method to scenarios with multiple exposures and apply it to the associations identified from the multi-exposure OPERA model (i.e., PPA >0.9). The basic idea behind HEIDI is that if the exposure-outcome association is driven by the same causal variant(s) because of either causality or horizontal pleiotropy, b_{xy} estimated at any xQTL SNP (denoted as xSNP) in LD with the causal variant(s) is expected to be identical to that estimated at the causal variant(s). The underlying causal variants are unknown, but this hypothesis can be tested by assessing the heterogeneity in $\widehat{b}_{xy}$ among all xSNPs in LD with the causal variant(s). To avoid performing all pairwise comparisons, the top xSNP is used as the target, and xSNPs in LD with the top xSNP are selected to test against the target.[16] To generalize the method to scenarios with multiple exposures, we define $\widehat{d}_{x_i y(s)} = \widehat{b}_{x_i y(s)} - \widehat{b}_{x_i y(top_i)}$, where the subscript "$i$" represents the i^{th}

molecular phenotype, *s* represents an instrument from a set of xSNPs present in all the xQTL datasets and passing an xQTL p value threshold, and the subscript "top_i" represents the lead xSNP for molecular phenotype *i*. The xSNPs are required to be instrument SNPs (i.e., $P_{xQTL} < 1.6 \times 10^{-3}$) for all tested molecular phenotypes. The null hypothesis of the multi-exposure HEIDI test is that all $d_{x_iy(s)} = 0$ and the alternative hypothesis of the multi-exposure HEIDI test is that any $d_{x_iy(s)} \neq 0$. If $\hat{d}_{x_iy(s)}$ is computed for two xSNPs *s* and *m*, then the covariance between $\hat{d}_{x_iy(s)}$ and $\hat{d}_{x_jy(m)}$ is

$$\mathrm{cov}\left(\hat{d}_{x_iy(s)}, \hat{d}_{x_jy(m)}\right) = \mathrm{cov}\left(\hat{b}_{x_iy(s)}, \hat{b}_{x_jy(m)}\right) - \mathrm{cov}\left(\hat{b}_{x_iy(s)}, \hat{b}_{x_jy\left(top_j\right)}\right) - \mathrm{cov}\left(\hat{b}_{x_iy(top_i)}, \hat{b}_{x_jy(m)}\right) + \mathrm{cov}\left(\hat{b}_{x_iy(top_i)}, \hat{b}_{x_jy\left(top_j\right)}\right)$$

where subscript *j* represents another molecular phenotype. For the ease of demonstration, we showed an example with only two xSNPs (i.e., *s* and *m*). It is worth noting that the multi-exposure HEIDI test is designed to include all the selected cis instrument SNPs. In the absence of sample overlap between the xQTL and GWAS data, the covariance between $\hat{b}_{x_iy(s)}$ and $\hat{b}_{x_jy(m)}$ is approximately

$$\mathrm{cov}\left(\hat{b}_{x_iy(s)}, \hat{b}_{x_jy(m)}\right) = \mathrm{cov}\left(\frac{\hat{b}_{z_sy}}{\hat{b}_{z_sx_i}}, \frac{\hat{b}_{z_my}}{\hat{b}_{z_mx_j}}\right) \approx \frac{\mathrm{cov}\left(\hat{b}_{z_sy}, \hat{b}_{z_my}\right)}{\hat{b}_{z_sx_i}\hat{b}_{z_mx_j}} + \frac{\mathrm{cov}\left(\hat{b}_{z_sx_i}, \hat{b}_{z_mx_j}\right) b_{z_sy} b_{z_my}}{\hat{b}^2_{z_sx_i}\hat{b}^2_{z_mx_j}}$$

where $\mathrm{cov}(\hat{b}_{z_sy}, \hat{b}_{z_my}) = r_{sm}\sqrt{\mathrm{var}(\hat{b}_{z_sy})\mathrm{var}(\hat{b}_{z_my})}$ with r_{sm} being the LD correlation between SNPs *s* and *m*, $\mathrm{cov}(\hat{b}_{z_sx_i}, \hat{b}_{z_mx_j}) = r_{sm}\rho_{x_ix_j}\sqrt{\mathrm{var}(\hat{b}_{z_sx_i})\mathrm{var}(\hat{b}_{z_mx_j})}$ with $\rho_{x_ix_j}$ being the between-study correlations due to sample overlap. We estimate $\rho_{x_ix_j}$ using the xQTL effects of null SNPs ($P_{xQTL} > 0.01$) by the Qi et al. approach[57] for each pair of tested exposure sites. When the pairwise xQTL datasets are from independent samples, $\rho_{x_ix_j} = 1$ when $i = j$ and 0 otherwise. At each locus where we observe association between multiple exposures (e.g., *t* exposures) and the outcome, we compute $\hat{d}_{x_iy(s)}$ for multiple xSNPs for each exposure and test if $d_{x_iy(s)} = 0$ for all the included xSNPs across *t* exposures by a multivariate test,[58] i.e., $\hat{\boldsymbol{d}} \sim \mathrm{MVN}(\boldsymbol{d}, \mathbf{V})$ with vector $\hat{\boldsymbol{d}} = (\hat{d}_{x_1y(s)} \cdots \hat{d}_{x_ty(s)} \cdots \hat{d}_{x_1y(m)} \cdots \hat{d}_{x_ty(m)})$ and variance-covariance matrix

$$\boldsymbol{V} = \begin{pmatrix} \mathrm{var}\left(\hat{d}_{x_1y(s)}\right) & \cdots & \mathrm{cov}\left(\hat{d}_{x_1y(s)}, \hat{d}_{x_ty(m)}\right) \\ \vdots & \ddots & \vdots \\ \mathrm{cov}\left(\hat{d}_{x_ty(s)}, \hat{d}_{x_1y(m)}\right) & \cdots & \mathrm{var}\left(\hat{d}_{x_ty(m)}\right) \end{pmatrix}$$

In our data analysis, we included in the multi-exposure HEIDI test only the SNPs present in the GWAS and all the xQTL datasets with $P_{xQTL} < 1.6 \times 10^{-3}$ (equivalent to $\chi^2 > 10$; as suggested by a previous study on Mendelian randomization to avoid weak instrumental variables[59]) for all the included exposures to avoid weak instrumental variables,[8] and pruned SNPs for LD with an r^2 threshold of 0.9. We used the ancestry-matched 1000 Genomes Project-imputed HRS data[60] as the reference sample to estimate the LD correlations between SNPs. When there is potential mismatch between datasets, we recommend applying methods such as DENTIST[61] to detect heterogeneity between datasets before running OPERA. Of note, the multi-exposure HEIDI can be applied to each individual exposure (identical to HEIDI in this case) and any combination of exposures. Different from the joint SMR analysis, the multi-exposure HEIDI test uses the marginal SMR effects because LD is canceled out between the marginal GWAS effect (the numerator) and the marginal xQTL effect (the denominator) under the null hypothesis, i.e.,

$$b_{x_iy(s)} = \frac{b_{z_sy}}{b_{z_sx_i}} = \frac{b_{z_cy} r_{sc}}{b_{z_cx_i} r_{sc}} = \frac{b_{z_cy}}{b_{z_cx_i}} = b_{x_iy(c)}$$

where *c* is the causal SNP, *s* is a tag SNP, and r_{sc} is the LD correlation between the two. It is worth noting that when the causal exposure is absent from the data, the cross-exposure leakage effect cannot be controlled using the joint SMR effects but can be detected by the multi-exposure HEIDI test.

Correcting for the cross-exposure leakage effect

One of the challenges to model the effects of multiple molecular phenotypes on a trait stems from LD between xQTL across molecular phenotypes, as illustrated in the following example (Figure S1A). Let us consider a locus with two SNPs, where SNPs 1 and 2 are the causal variants for exposures 1 and 2, respectively, and only exposure 1 affects the outcome. In a single-exposure analysis such as SMR, exposure 2 would appear to be associated with the outcome if the two SNPs are in LD (Figure S1A). This is because the effect of exposure 2 on the outcome is estimated by the ratio of the effect of SNP 2 on exposure 2 to that on the outcome, and the effect of SNP 2 on the outcome is non-zero owing to the LD between the two SNPs, which seems that the effect of exposure 1 on the outcome leaks to exposure 2 (Figure S2A). Hence, in the presence of the cross-exposure leakage effect, using marginal exposure effects (e.g., those estimated from SMR) to compute the likelihoods in OPERA would lead to erroneous posterior inference about the configurations and thereby false discoveries of the exposure-outcome associations (Figure S7). It is of note that although applying HEIDI test in this scenario can reject the associations caused by linkage, the $\boldsymbol{\pi}$ estimates for alternative configurations would be downward biased due to selection based on the HEIDI test.

Therefore, we propose to use the joint SMR effect in OPERA analysis. As shown in Figure S1A, if the xQTL for a null exposure is in LD with the xQTL for the causal exposure, the estimated marginal effect for the null exposure, $\widehat{b}_{x_iy} = \widehat{b}_{z_iy}/\widehat{b}_{z_ix_i}$, could be significantly different from zero, because b_{z_iy} captures the causal effect through LD (Figure S2A), referred to as cross-exposure leakage effect. To address this issue, we propose to use the joint SMR effect in OPERA, which is computed as the ratio of the joint GWAS effect of the lead xSNP over its marginal xQTL effect. The joint GWAS effect of the lead xSNP (z_i) for exposure i, β_{z_iy}, is estimated by fitting it jointly with the COJO[46] xSNPs for the other exposures ($\beta_{z_{-i}y}$), where $\mathbf{z}_{-i}$ represent COJO xSNPs except for those for exposure i after LD pruning with an r^2 threshold of 0.9. The logic is that if $\widehat{\beta}_{z_iy}$ shrinks to zero conditional on the effects of xSNPs for the other exposures, it is unlikely that exposure i has a causal effect on the outcome. Following the summary data-based joint analysis method in COJO,[46] we estimate the joint GWAS effects for the selected xQTL, $\widehat{\beta}_{zy} = [\widehat{\beta}_{z_iy}, \widehat{\beta}_{z_{-i}y}]'$, as

$$\widehat{\beta}_{zy} = \mathbf{W}^{-1}\mathbf{D}\widehat{\boldsymbol{b}}_{zy} \text{ and } var(\widehat{\boldsymbol{\beta}}_{\boldsymbol{zy}}) = \sigma^2_{\mathrm{J}}\mathbf{W}^{-1}$$

where $\widehat{\boldsymbol{b}}_{zy}$ is a vector of the marginal GWAS effects, $\mathbf{W}$ is the variance-covariance matrix of SNP genotypes computed from a reference panel and $\mathbf{D}$ is a diagonal matrix of $\mathbf{W}$, and σ^2_{J} is the residual variance in the joint analysis. We can estimate σ^2_{J} by

$$\widehat{\sigma}^2_{\mathrm{J}} = \frac{\boldsymbol{y}'\boldsymbol{y} - \widehat{\boldsymbol{\beta}}'_{zy}\mathbf{D}\widehat{\boldsymbol{b}}_{zy}}{(n-p)}$$

where n is the GWAS sample size, p is the number of selected xQTL, and $\boldsymbol{y}'\boldsymbol{y}$ is estimated by the median value of $D_i(SE_i^2(n-1)+\widehat{b}_i)$ of the selected xQTL, with SE_i and $\widehat{b}_i$ being the standard error and estimated effect size from the GWAS summary statistics. We show in Figure S30 that $\widehat{\sigma}^2_{\mathrm{J}}$ estimated from the subset of SNPs in common between GWAS and *cis*-xQTL is highly consistent with that estimated from the full set of GWAS SNPs. Since the xQTL effect for the target exposure is independent of those for the other exposures, i.e., $\widehat{b}_{z_ix_i} = \widehat{b}_{z_ix_i}|\boldsymbol{x}_{-i}, \widehat{\beta}_{x_{-i}y}$, we do not need to adjust the marginal xQTL effect (See Data S2 for full derivation). The so-called joint SMR effect for exposure i on outcome is then estimated by

$$\widehat{\beta}_{x_iy} = \frac{\widehat{\beta}_{z_iy}}{\widehat{b}_{z_ix_i}}$$

Simulations based on imputed genotype data from the UK Biobank

We performed the following simulations based on the imputed genotype data from the UK Biobank after quality controls (QC).[29] We included 1,164,362 HapMap3 SNPs with MAF >0.01, Hardy-Weinberg equilibrium test $p > 1 \times 10^{-6}$, genotyping rate >0.95, and imputation information score >0.8, and selected 348,501 unrelated individuals, with a relatedness threshold of 0.05, based on the genetic relatedness estimated from the SNP set above using GCTA.[62]

We first performed simulations to assess the performance of OPERA in estimating $\boldsymbol{\pi}$, which required genome-wide independent loci. We randomly sampled 400 approximately independent LD blocks with size >500Kb in the genome defined by Pickrell et al.[63] In each LD block (l), we randomly sampled SNPs within a 500 kb window (number of variants ranging from 31 to 758 with a median of 212), among which m_{li} SNPs were chosen at random to be the causal variants for each exposure i. We assumed $m_{li} \sim$ Poisson (2). The phenotype measures of exposure i at locus/block l (i.e., $\boldsymbol{x}_{li}$) were simulated by

$$\boldsymbol{x}_{li} = \mathbf{Z}_{li}\boldsymbol{b}_{z_{li}x_{li}} + \boldsymbol{e}_{li}$$

where $\mathbf{Z}_{li}$ is the genotype matrix for the causal variants, $\boldsymbol{b}_{z_{li}x_{li}} \sim N(\mathbf{0}, \mathbf{I})$ and $\boldsymbol{e}_{li} \sim N(0, var(\mathbf{Z}_{li}\boldsymbol{b}_{z_{li}x_{li}})/(1/h^2_x - 1))$ with $h^2_x = 0.1$ being the proportion of variance in the exposure explained by the causal variants. We standardised $\boldsymbol{x}_{li}$ to have mean 0 and variance 1. We simulated the outcome by an additive model for the exposure effects across t exposures and m genomic loci,

$$\boldsymbol{y} = \sum_{l=1}^{m}\sum_{i=1}^{t}\boldsymbol{x}_{li}b_{x_{li}y} + \boldsymbol{e}_y$$

where $b_{x_iy} \sim \pi_i N(0, \sigma^2_{b_i}) + (1-\pi_i)\varphi$ and $\boldsymbol{e}_y \sim N(0, var(\sum_{l=1}^{m}\sum_{i=1}^{t}\boldsymbol{x}_{li}b_{x_iy})/(1/R^2_{xy} - 1))$ with $R^2_{xy} = 0.6$ being the proportion of variance in the outcome explained by all the exposures. We set $m = 400$ and $t = 3$ for simulations based on the UK Biobank data (see Figure S5 for simulation result when $t = 5$ with unlinked markers). We first randomly sampled loci with 1, 2 or 3 causal exposures and assigned the sampled causal effects to the exposures, which assumes a non-independence structure in $b_{x_{li}y}$. The proportion of loci with 0, 1, 2 or 3 exposures associated with the outcome was set to be 0.78, 0.05, 0.02 or 0.01, constituting the true $\boldsymbol{\pi}$ values. Under this setting, the proportion of causal sites for each exposure was 0.1, and each causal site explained 0.5% of the variance in the outcome on average. We sampled the causal effects from $N(0, \sigma^2_{b_i})$ with $\sigma^2_{b_i} = 0.02$ for all exposures, approximately equal to that from the pairwise SMR analyses in real data. Given the simulated outcome and exposures data, we performed a standard GWAS with a subset of 300,000 individuals and xQTL analyses using independent samples with a range of sample sizes (from 500 to 10,000). We then applied OPERA to estimate $\boldsymbol{\pi}$ through the stage-1 analysis. The whole simulation process was repeated 100 times.

We used the estimated π to perform the stage-2 OPERA analysis, and evaluated the power, FPR and FDR of OPERA across simulation replicates. The power was quantified as, among loci with exposure *i* (or a combination of exposures $\boldsymbol{c} = i,\ldots,j$) being causal, the proportion of which had the marginal PPA_i (or the joint $PPA_{\boldsymbol{c}}$) greater than a cut-off value (e.g., 0.9) and $P_{HEIDI} > 0.01$. The FPR was computed as, among loci with exposure *i* (or $\boldsymbol{c}$) being null, the proportion of which had the marginal PPA_i (or the joint $PPA_{\boldsymbol{c}}$) greater than the threshold. The FDR was computed as, among loci where the marginal PPA_i for exposure *i* (or the joint $PPA_{\boldsymbol{c}}$ for exposures $\boldsymbol{c}$) was greater than the threshold, the proportion of which with the exposure *i* (or $\boldsymbol{c}$) being null. Following Equations 5 and 6, we computed the estimated FDR and FPR at a range of PPA thresholds and compared them with the observed FDR and FPR from the simulation. To investigate the relationship between PPA and true discovery rate, we stratified the reported PPA into 10 even bins and then quantified the true discovery rate in each PPA bin.

We checked the performance of OPERA under unbalanced design by randomly setting 10% of the null sites as missing for each exposure. We performed the stage-1 and the stage-2 analysis based on this setting and further quantified the power, FPR and FDR as above. We also checked the performance of OPERA under the scenario with multiple sites per exposure by simulating an additional null site at each locus. We performed the stage-2 analysis and multiple-exposure HEIDI test for all possible combinations of sites across exposures and then quantified the power, FPR and FDR at a given PPA threshold along with a HEIDI threshold of 0.01. To compare the power between the joint-exposure analysis and the single-exposure analysis using OPERA to detect causal exposures, we simulated a dataset with 3 exposures based on the unbalanced setting with the proportion of variance in the outcome explained by each site being $R^2_{x_1y} = 0.01\%$, $R^2_{x_2y} = 0.1\%$ and $R^2_{x_3y} = 1\%$ for exposures 1, 2 and 3, respectively.

To calibrate the multi-exposure HEIDI test, we first simulated a homogeneous causal model (i.e., the null model for the HEIDI test), where the two causal exposures and the outcome shared the same set of causal variants (Figures S1B and S1C). We then simulated a heterogeneous causal model (i.e., the alternative model for the HEIDI test), where only the first exposure affected the outcome, and the causal variant of the second exposure was in LD ($0.1 < r^2 < 0.9$) with that of first exposure (Figure S1A). Finally, we simulated another heterogeneous causal model, where both exposures affected the outcome and the two causal variants for two corresponding exposures were in LD ($0.1 < r^2 < 0.9$; Figure S1D). We repeated the simulation process 10,000 times for each model and generated the quantile-quantile plot based on the reported multi-exposure HEIDI test p values (Figure S10).

Data used in this study

We used seven types of xQTL data in our analysis. The summary-level eQTL data were from the eQTLGen data ($n = 31,684$),[22] which was a meta-analysis of blood samples comprising ~9 million SNPs and 19,250 genes. We also used the blood eQTL summary statistics from the CAGE data ($n = 2,765$), which comprised 38,624 normalized gene expression probes and ~8 million SNPs. The caQTL summary statistics were from lymphoblastoid cell lines samples of British ancestry (n = 100).[24] There were 277,128 chromatin accessibility peaks on autosomes measured by assay for transposase-accessible chromatin using sequencing (ATAC-seq), 13,949 of which had at least one caQTL at $p < 5 \times 10^{-8}$. The peripheral blood mQTL summary data were from a meta-analysis of two blood mQTL datasets from McRae et al.[16,26] ($n = 1,980$). The methylation states were measured based on Illumina HumanMethylation450 chips and 94,338 methylation probes were identified with at least one *cis*-mQTL. For the BLUEPRINT hQTL summary data, two histone modification marks (i.e., H3K27ac and H3K4me1) were measured in three major immune cell types ($CD14^+$ monocytes, $CD16^+$ neutrophils, and $CD4^+$ naive T cells) in up to 197 individuals.[25] We chose hQTL data from the $CD14^+$ monocytes because it provided the largest number ($m = 18,153$) of histone marks with at least one hQTL with $p < 5 \times 10^{-8}$. The summary statistics of SNPs on plasma proteins (i.e., pQTL) were from the INTERVAL dataset[23] (n = 3,301). It comprised of 3,622 plasma proteins or protein complexes assayed using 4,034 modified aptamers on 3,301 participants. There were 698 proteins with at least one *cis*-pQTL. The apaQTL summary data were from the a recent 3′UTR alternative polyadenylation analysis using the RNA-seq data from GTEx blood tissue.[27] The APA events were detected by DaPars v.2.0 algorithm and there are 1,050 APA events with at least one *cis*-apaQTL. The blood sQTL data were from the GTEx dataset. The RNA splicing event were quantified by leafcutter and there were 6,638 splicing events with at least one *cis*-sQTL. We included in our analysis only the molecular phenotype sites with at least one *cis*-xQTL at $P_{xQTL} < 5 \times 10^{-8}$ and excluded the molecular phenotype sites in the major histocompatibility complex (MHC) region because of the complexity of this region.

We included in this study 19 complex traits from the latest available GWAS meta-analyses and 31 complex traits from the UK Biobank. They are height,[33] BMI,[33] high-density lipoprotein,[37] low-density lipoprotein,[37] thyroglobulin,[37] educational years,[41] rheumatoid arthritis,[38] schizophrenia,[40] coronary artery disease,[36] type 2 diabetes,[35] Crohn disease,[38] ulcerative colitis,[38] Alzheimer's disease,[39] breast cancer,[43] prostate cancer,[43] high cholesterol,[29] monocyte count,[34] neutrophil count,[34] red blood count[34] and 31 complex traits from the UK Biobank included in the Zeng et al.[64] We used acronyms to label complex traits as showed in Figure 5. They are body mass index (BMI), high-density lipoprotein (HDL), low-density lipoprotein (LDL), thyroglobulin (TG), educational years (EY), rheumatoid arthritis (RA), schizophrenia (SCZ), coronary artery disease (CAD), type 2 diabetes (T2D), Crohn disease (CD), ulcerative colitis (UC), Alzheimer's disease (ALD), breast cancer (BrCancer), prostate cancer (PrCancer), high cholesterol (HC), monocyte cell count (MonoCC), neutrophil cell count (NeutCC), red blood cell count (RedBCC), basal metabolic rate (BMR), body fat percentage (BFP), hand grip strength left (HGSL), hand grip strength right (HGSR), hip circumference adjusted for BMI (HC), Hell BMD T-score (HBMD), waist circumference adjusted for BMI (WC), weight (WT), waist–hip ratio adjusted for BMI (WHR), diastolic blood pressure

(DBP), forced expiratory volume (FEV), forced vital capacity (FVC), peak expiratory flow (PEF), pulse rate (PR), systolic blood pressure (SBP), fluid intelligence score (FIS), mean time to correctly identify matches (MTCIM), neuroticism score (NS), age at first live birth (AFLB), age at menopause (Mnps), Age menarche (Mnrch), birth weight (BW), male pattern baldness (MPB), Allergic rhinitis (AR), Dyslipidemia (Dyslp), Hemorrhoids (Hmrr), Hypertensive disease (HD), varicose veins (VV). The SNP effects on quantitative traits were in standard deviation units, and those on disease traits (e.g., case-control design) were expressed as log odds-ratios. Only common variants with a minor allele frequency >0.01 were retained for analysis. We used the 1000 Genomes Project (1KGP) imputed data of the Health and Retirement Study (HRS)[60] (n = 8,557) as the LD reference data required for the joint SMR analysis and multi-exposure HEIDI test.